Technical Report 1047

Research Methods for Advanced Warfighting Experiments

Carl W. Lickteig
U.S. Army Research Institute

August 1996

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

David W. Bessemer
Scott E. Graham
William M. Parry
Jeffery R. Witsken

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1996, September	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) January 1995-July 1996
4. TITLE AND SUBTITLE Research Methods for Advanced Warfighting Experiments		5a. CONTRACT OR GRANT NUMBER
		5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) Carl W. Lickteig		5c. PROJECT NUMBER A791
		5d. TASK NUMBER 2228
		5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RZ 5001 Eisenhower Avenue Alexandria, VA 22333-5600		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Technical Report 1047

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

Advanced Warfighting Experiments (AWEs) typify the Army's emerging need for more pragmatic and responsive research methods to address the changing climate of military research and improve future force capability. Formative force improvement enables or mediates the summative objective--a more capable force. To help achieve the primary objective, this report recommends the AWEs adapt formative evaluation methods that focus on exploration, explanation, and improvement. This report identifies a set of key fundamental and formative method issues for the AWEs and provides corresponding method recommendations for more reliable and useful AWE findings. The report's method recommendations embed a mechanism of expanded AWE evaluation teams that implement lessons learned into "living products" for Army-wide Force XXI efforts.

15. SUBJECT TERMS

Advanced Warfighting Experiments   Human performance   Formative evaluation
TTPs   Research methods   Digital systems

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	78	

Technical Report 1047

Research Methods for Advanced Warfighting Experiments

Carl W. Lickteig
U.S. Army Research Institute

Armored Forces Research Unit
Barbara A. Black, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

August 1996

Army Project Number **Education and Training Technology**
20262785A791

FOREWORD

As part of the U.S. Army Research Institute for the Behavioral and Social Sciences' (ARI) program to train the force, the objective of the Future Battlefield Conditions (FBC) team at Fort Knox is to enhance soldier preparedness through development of training and evaluation methods to meet future battlefield conditions. The FBC's work is performed under Work Package 2228, FASTTRAIN, Force XXI Training Methods and Strategies. ARI's research on training requirements and evaluation methods is supported by a Memorandum of Agreement between the U.S. Army Armor Center (USAARMC) and ARI titled Manpower, personnel and Training Research, Development, Test, and Evaluation for the Mounted Forces, 16 October 1995.

The U.S. Army is embarked on a venture into the 21st century with a modernization effort called Force XXI. Supporting Army research efforts are focused by a challenging series of ongoing Advanced Warfighting Experiments (AWEs). Formative force improvement enables or mediates the summative objective--a more capable force. To help achieve the primary objective, this report recommends the AWEs adapt formative evaluation methods that focus on exploration, explanation, and improvement. This report identifies key fundamental and formative method issues for the AWEs and provides corresponding method recommendations for more reliable and useful AWE findings.

The findings of this report may help clarify the AWE's purpose and related expectations. The method recommendations provide useful guidance to AWE evaluators concerning the conduct of AWEs. These method recommendations embed a mechanism of expanded AWE evaluation teams that implement lessons learned into living products for Army-wide Force XXI efforts.

ZITA M. SIMUTIS
Deputy Director
(Science and Technology)

EDGAR M. JOHNSON
Director

RESEARCH METHODS FOR ADVANCED WARFIGHTING EXPERIMENTS

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army is embarked on a venture into the 21st century with a modernization effort called Force XXI. The Army's research in support of this challenging effort is focused by an ongoing series of Advanced Warfighting Experiments (AWEs). This report addresses how military researchers might more effectively apply research methods to the AWEs.

Procedure:

The AWEs typify the Army's emerging need for more pragmatic and responsive research methods to address the changing climate of military research and improve future force capability. Formative force improvement enables or mediates the summative objective—a more capable force. To help achieve the primary objective, this report recommends the AWEs adapt formative evaluation methods that focus on exploration, explanation, and improvement.

Findings:

This report identifies a set of key fundamental and formative method issues for the AWEs and provides corresponding method recommendations for more reliable and useful AWE findings. To exemplify the use of AWE formative research methods, the report focuses on two enabling objectives: the Tactics, Techniques, and Procedures (TTPs) to exploit information technologies; and the process required to provide a relevant, common picture of the battlefield to Force XXI combatants and supporters. The methods proposed focus on implementing findings into "living products."

Utilization of Findings:

The findings of this report may help clarify the AWE's purpose and related expectations. The method recommendations provide useful guidance to AWE evaluators concerning the conduct of AWEs. These method recommendations embed a mechanism of expanded AWE evaluation teams that implement lessons learned into living products for Army-wide Force XXI efforts.

RESEARCH METHODS FOR ADVANCED WARFIGHTING EXPERIMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

RESEARCH METHODS FOR ADVANCED WARFIGHTING EXPERIMENTS

Introduction

The U.S. Army is embarked on a venture into the 21st century. Joint Venture, in fact, is the name of the Army's main effort in its campaign plan toward force modernization. The plan for Joint Venture is to aggressively execute an iterative cycle of concept development, force design, and experimentation to achieve this modernization objective, Army XXI (U.S. Department of the Army, 1995a). This report focuses on the Army's Force XXI Advanced Warfighting Experiments (AWEs) supporting this objective, particularly research methods appropriate to the AWEs and related Army modernization efforts.

The broad scope, purpose and agenda of the AWEs pose a serious challenge to military researchers and more traditional approaches to military testing. To meet this challenge, the Army's research community might adapt and devise research methods appropriate to the AWEs, and assist military leaders in AWE design, conduct and interpretation. The AWEs typify the Army's emerging need for more pragmatic and responsive research methods to address the changing climate of military research (Hollis, 1995; O'Bryon, 1995). This Introduction section reviews the AWE research plan and methods relative to Joint Venture objectives, and then considers some of the research challenges inherent to contemporary military research and the AWEs.

Research objectives should determine research methods. A basic premise of this report is that the primary objective of Joint Venture and the AWEs is improving future force capability. A corollary premise is that a subsequent objective of these same efforts is "proving" improved force capability. Formative force improvement enables or mediates the summative objective, a more capable force. Differential methods and measures are required for these different objectives. This report, therefore, focuses on research methods and measures directed at improving the force and applicable to the AWEs and related Army-wide efforts. The rationale for this approach is that if intermediate objectives are not met, more final objectives may not be attained.

This report suggests that the AWEs employ an overarching research strategy such as program evaluation that includes both formative and summative evaluations (Scriven, 1967). The

expansive research methods of program evaluation may effectively encompass the broad scope of AWE objectives, including improving and proving force capability. While formative and summative evaluations can be conducted concurrently in the AWEs, a preliminary concern with formative issues and methods may avoid summative conclusions of failure.

Formative and summative evaluations are complementary and legitimate, scientifically acceptable, methods for conducting research that entail many of the same fundamental research methods, as discussed in this report. Formative and summative evaluations differ primarily in their focus and role. Formative evaluations focus on intermediate goals and play a productive role in their attainment; summative evaluations address more terminal goals and adjudicate their attainment.

The more responsive methods of formative evaluation are well suited to the broad scope, demanding pace, and macro-level complexity of the AWE's. The exploratory and explanatory power of formative evaluation are needed to inform the design of Force XXI. The more exacting methods of summative evaluation are not precluded in the AWEs, but are best enforced in subexperiments. True soldier-in-the-loop subexperiments require the more restrictive methods of the scientific experiment, not mere miniaturizations of an AWE's macro complexity.

The Method section focuses on key fundamental and formative evaluation methods for the AWEs. It identifies twelve AWE method issues and provides corresponding recommendations to address each of the research issues raised. Eight of these twelve issues are regarded as fundamental, common to all evaluation. The method recommendations presented herein for these fundamental evaluation issues are appropriate to all AWEs, formative or otherwise.

The four remaining method issues considered in the Method section are more unique to the conduct of AWE formative evaluations. The method recommendations for these issues are especially appropriate for attaining some of the key intermediate objectives of Force XXI. Two intermediate objectives are used to exemplify formative research methods: the Tactics, Techniques, and Procedures (TTPs) to exploit information technologies; and, the process required to provide a relevant, common picture of the battlefield to Force XXI combatants and supporters.

An acceptable exit criterion for the AWEs, as formative evaluations, may be "living products" that implement lessons learned during and between AWEs. These products should span the domains of Doctrine, Training, Leadership, Organization, Materiel and Soldiers (DTLOMS). Additional exit criteria are expected from the AWEs and Joint Venture (U.S. General Accounting Office, 1995). Sponsors and decision makers await summative and defensible conclusions that quantify the force-level benefits and costs associated with advanced information systems. This report's formative focus, however, suggests that hypothesized Force XXI capabilities should be more fully developed before they are summarily tested.

The AWEs are not stand-alone evaluations. The methods presented here stress that an AWE's core evaluation team could be strongly buttressed by expanded teams drawn from related Army Research, Development, Acquisition and Training (RDA&Tng) programs. For summative objectives, these RDA&Tng programs and the Army's Battle Labs provide an appropriate forum for conducting scientific experiments that contribute to the body of evidence supporting the AWEs and Force XXI. For formative objectives, members of these expanded teams might ensure that lessons learned on advanced information systems from their ongoing programs and the AWEs, are iteratively implanted in a common set of living products. This report's AWE method recommendations embed a mechanism for sustaining and employing these products across AWEs and Army-wide Force XXI efforts.

Joint Venture's Process of Change

As envisioned by the Army's senior leadership, "Force XXI is a comprehensive approach to redesign the force--organized around information--to be inherently more versatile and flexible. ....Force XXI is about creating the world's best Army for the 21st century" (Sullivan, 1995). The Joint Venture Campaign Plan, by Training and Doctrine Command (TRADOC), describes the steps and responsibilities for the Army's central axis of effort, see Figure 1, in achieving the operational force of the future (U.S. Department of the Army, 1995a). That plan is based on iterative assessments, particularly the AWEs, of how the future Army should equip, train and fight the force. The AWE's pervasive nature includes the domains of DTLOMS across all operations, echelons and operating systems (U.S. Department of the Army, 1995f).

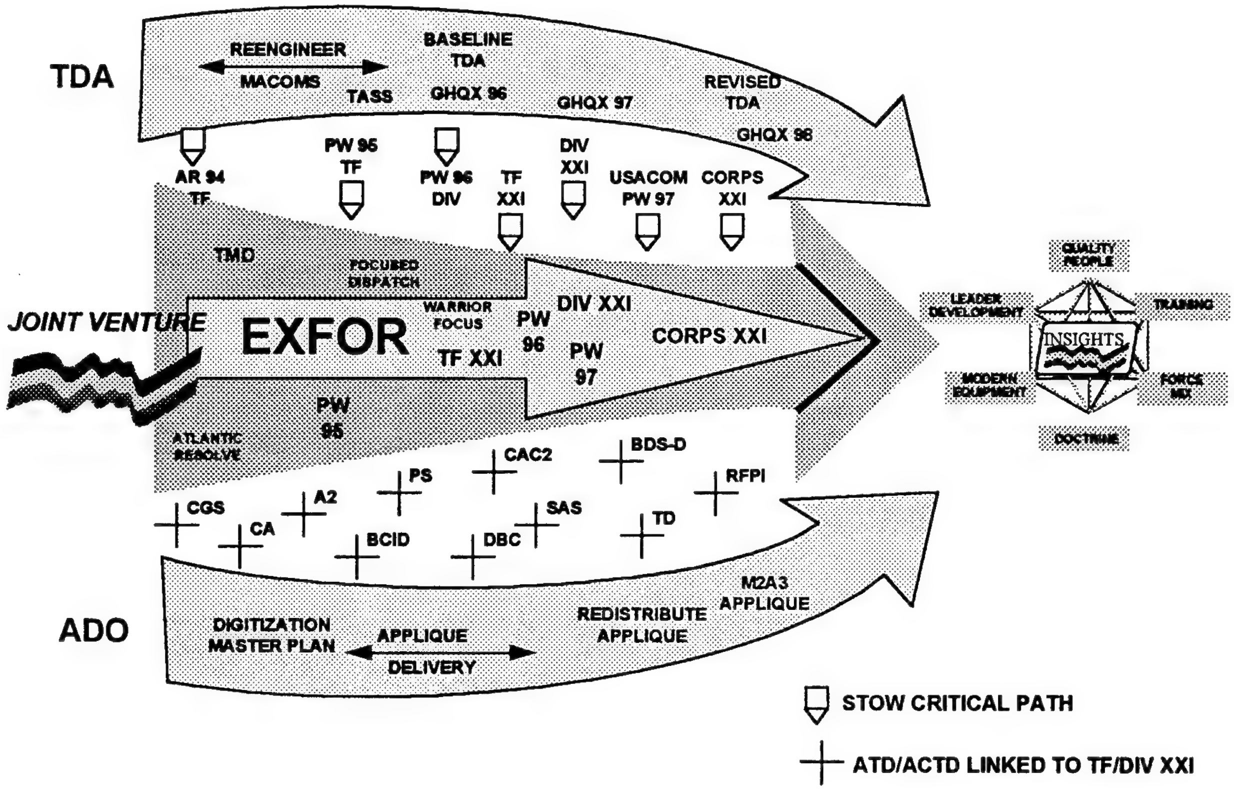

Figure 1. The Force XXI Campaign Plan with Joint Venture as main axis (Adapted from U.S. Army Armor Center, 1995c).

One of the most fundamental characteristics of this Army modernization effort is process, a process of continuous transformation. Army leadership has stressed that the development of Force XXI involves continuous experimentation, discovery learning, and iterative refinement. This open-minded approach seems particularly appropriate in view of the digital information technologies and capabilities that must enable many of the Army's future development, training, and evaluation efforts. Exponential increases in the ability of digital systems to process and globally access and distribute information, make even visionaries reluctant to define endpoints or outcomes (U.S. Department of the Army, 1994b).

Key objectives directing the Joint Venture effort are to establish deliberate Force XXI patterns of operations: project the force, protect the force, gain information dominance, shape the battlespace, decisive operations, and sustain the force (U.S. Department of the Army, 1996a). Clearly, advanced information technologies are key to establishing these patterns. The potential power of information systems as a force multiplier is

escalating information to the level of a battlefield weapon system. A salient indicator of the perceived importance of information is the Army's initial publication of Field Manual 100-6, Information Operations (U.S. Department of the Army, 1995d). Similarly, the Army's emerging doctrine is predicated on the ability of information technologies to provide unprecedented military capability, such as a common and relevant depiction of the battlefield situation to all force combatants and supporters (U.S. Department of the Army, 1995e).

The requirement to focus on process and iterative products versus end-state products, holds for Joint Venture's training and force development efforts. On the training side, a striking example of this emphasis on the process of improvement is the position of the Force XXI training development community, that future training programs and literature must be "living documents" (U.S. Department of the Army, 1995c). Accordingly, the Force XXI Training Program serves as a prototype of how emerging technologies and methods may be synthesized to improve military capability in the 21st century (Martin, 1995; Quinkert and Black, 1994). Training development leadership has stressed that training should readily adapt to changes and lessons learned across the Army, including lessons from the Combat Training Centers (CTCs) and the AWEs. Notably, changes in any DTLOMS domain often require related changes in other domains, such as Materiel changes that lead to changes in Training or Doctrine.

On the evaluation side, the process of achieving Force XXI capabilities is at issue. To exemplify research methods for addressing this issue, this report's method recommendations focus on achieving a common picture of the battlefield and the TTPs for employing information systems. Ideally, digital information systems are expected to provide a common, relevant picture of the battlefield scaled to the specific level of interest and needs of Force XXI combatants and supporters (U.S. Department of the Army, 1996a). Although not well defined, this common picture capability is a key intermediate product anticipated from Force XXI's advanced information systems. Expectations of improved force capability presume that future combatants will exploit this intermediate-level capability to achieve end-level improvements. Improvements that might be documented, for example, by force-level measures of effectiveness (MOEs).

A premature focus on AWE outcome measures, such as MOEs, might conclude that advanced information technologies do not

result in force improvement. A focus on intermediate measures, however, might conclude the common picture was never adequately achieved, or tested. An evaluative focus on the process required to attain and maintain the common picture, should precisely identify correctable process deficiencies, by source and type. Implementing, versus documenting, lessons learned about such deficiencies should directly support the intermediate objectives that lead to hypothesized Force XXI capabilities.

Advanced Warfighting Experiments

The AWEs are the principal activities in the Joint Venture plan which includes various types and levels of warfighting experiments (U.S. Department of the Army, 1996a). The Army's currently planned series of AWEs is envisioned as the focal effort, the central axis in Figure 1, in establishing Force XXI capabilities. By design, the AWEs are devised as macro-level evaluations, explorations of complex and interrelated issues such as force organization, doctrine, and the TTPs required for future Army operations. The AWEs are based on an iterative sequence and mix of warfighting simulations--live, constructive and virtual--in which soldiers and units conduct realistic tactical operations. These various simulations, or research settings, merge for some AWEs. For example, a hybrid live/virtual mix during the Focused Dispatch AWE linked members of the live Task Force on actual terrain, at the Western Kentucky Training Area in Greenville, with other members of the unit operating in virtual simulation (U.S. Army Armor Center (1995a). Future AWEs may be based on Synthetic Theater of War (STOW) technologies that will interactively link live, constructive, and virtual simulations (Cosby, 1995; Sottilare, 1995).

The AWEs are not the only military research and testing activities in support of Force XXI, as illustrated in Figure 1. The Army's Battle Labs support the AWEs and conduct their own warfighting experiments, Battle Lab Warfighting Experiments (BLWEs). These BLWEs are typically smaller, more focused assessments addressing a single battle dynamic, and may range from practical to scientific experiments depending upon the priorities and resources of the Army. The Joint Venture plan also includes the research and development efforts being performed under programs such as the Advanced Technology Demonstrations(ATDs) and Advanced Concept and Technology Demonstrations (ACTDs).

Conceptually, all of the Army's warfighting experiments are linked by a consistent set of hypotheses and experimental objectives (U.S. Department of the Army, 1995b). However, the scope and resources available to each of the proposed warfighting experiments may constrain the range of research issues examined. Underlying commonalities, such as a common set of MOEs and associated measures of performance (MOPs), are expected to track progress in achieving Force XXI goals and objectives. The Rolling Baseline assessment strategy for Joint Venture, developed by the Operational Test and Evaluation Command (OPTEC), is built from a generic set of MOEs and MOPs common to all warfighting experiments. This baseline is expected to document the current status of force effectiveness and trends of improvement across AWEs (U.S. Department of the Army, 1995b).

All of the Army's warfighting experiments begin with a formal hypothesis derived from selected DTLOMS issues or from the more macro concepts underlying Force XXI. The linchpin to these evaluation efforts is the fundamental hypothesis of the Joint Venture Campaign Plan:

> If we know the performance of a baseline organization, then we can apply digital technology to the organization, conduct experiments, and gain insights into improved battlefield performance which will enable us to redesign operational concepts and units to optimize military capabilities. (U.S. Department of the Army, 1995b)

This hypothesis reflects the iterative nature of the Force XXI efforts directed at improving the future force and Joint Venture's formative objectives.

Useful findings and applications could be garnered across all related Army efforts--AWEs, BLWEs, ATDs and ACTDs--for Force XXI's comprehensive reorganization and redesign. Methods that effectively compile and implement this information are a primary challenge that Joint Venture poses for the military research community, and a focus of this report's Method section.

<u>Current AWE Research Methods</u>

Consistent with this iterative approach, the analytic method for Force XXI is based on progressive cycles: Model-Experiment-Model-Validate (MEMV) (U.S. Department of the Army, 1995a,

1995b). This MEMV method is to address all forms of simulation included in the plan for Joint Venture, see Figure 2. The method attempts to leverage virtual and constructive simulations to cycle more rapidly through progressive iterations (U.S. Department of the Army, 1996a). Models used in this MEMV method are initially developed in concert with the Army's Battle Labs. Experiments in constructive or virtual simulation based on these models are expected to provide: insights to the DTLOMS; data to calibrate or improve the simulations being used; and, data for a rolling baseline. Modeling refinements use the data and insights from these experiments to recalibrate the force-on-force simulations. Validation is ideally slated in live simulation for all AWE models, such as refinements in doctrine, organization or training (U.S. Department of the Army, 1995a).

The Joint Venture Campaign Plan appears to be based, at least in part, on the model of a scientific experiment (Campbell & Stanley, 1963). Controlled experimentation is suggested by statements of formal hypotheses, the actual labeling of the AWEs as "experiments," a heavy reliance on outcome measures, and an emphasis on validating models of hypothesized improvement against a rolling baseline. Briefings on the AWE Baseline Assessment Strategy, for example, state "the scientific method paradigm must be adapted to the context of the AWEs" (Dubin, 1995). The AWE's macro-level complexity and developmental agenda, however, complicate imposing the exacting methods of the scientific experiment.

The current MEMV method is explained as a spiral development process in which validation establishes a new baseline for new objectives and a new round of MEMV in the same or subsequent AWEs. Figure 3 shows how Focused Dispatch tried to apply the spiral process and rolling baseline across simulation settings. Notably, the Rolling Baseline Strategy is not a simple, agreed upon concept. While the spiral development emphasis appears consistent with the formative nature of Joint Venture and the AWEs, the emphasis on a rolling baseline and validation may prematurely divert AWE efforts toward summative issues and methods.

Contemporary Challenges in Military Research

Some members of the military research community are calling for more responsive methods and approaches to the demanding nature of contemporary military research. In part, this may be a

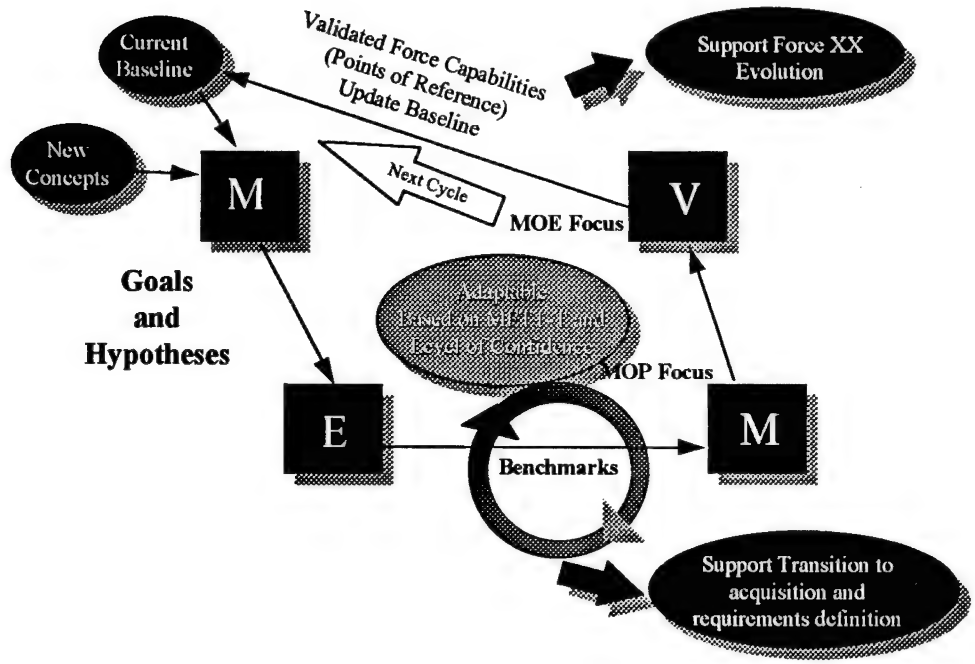

Figure 2. Model, Experiment, Model, Validate (MEMV) Methodology and the Rolling Baseline (Adapted from Dubin, 1995).

reaction to current criticisms that traditional validation methods are too slow to support the development and fielding of new technologies, given the Army's streamlined model for research, development and acquisition (Colwell, 1993). It may also be part of an emerging consensus that traditional test and evaluation methods used by military researchers are heavily biased toward identifying failures rather than evolving workable solutions to the military's problems (O'Bryon, 1995).

The pragmatic challenge to devise research methods for solving practical problems is not new or limited to members of the military's research community. Serious criticisms of nonmilitary researchers' reluctance to develop methods for exploratory and formative research issues are recurrent (McCall & Bobko, 1990). The need for more responsive and formative research issues was ably articulated by Bouchard (1976):

> This is not to imply that scientific knowledge and rigorous procedures should not be used when they are applicable..., but rather to emphasize that the context of discovery had hardly been mined while the context of justification had been overburdened with trivial investigations. (p. 366)

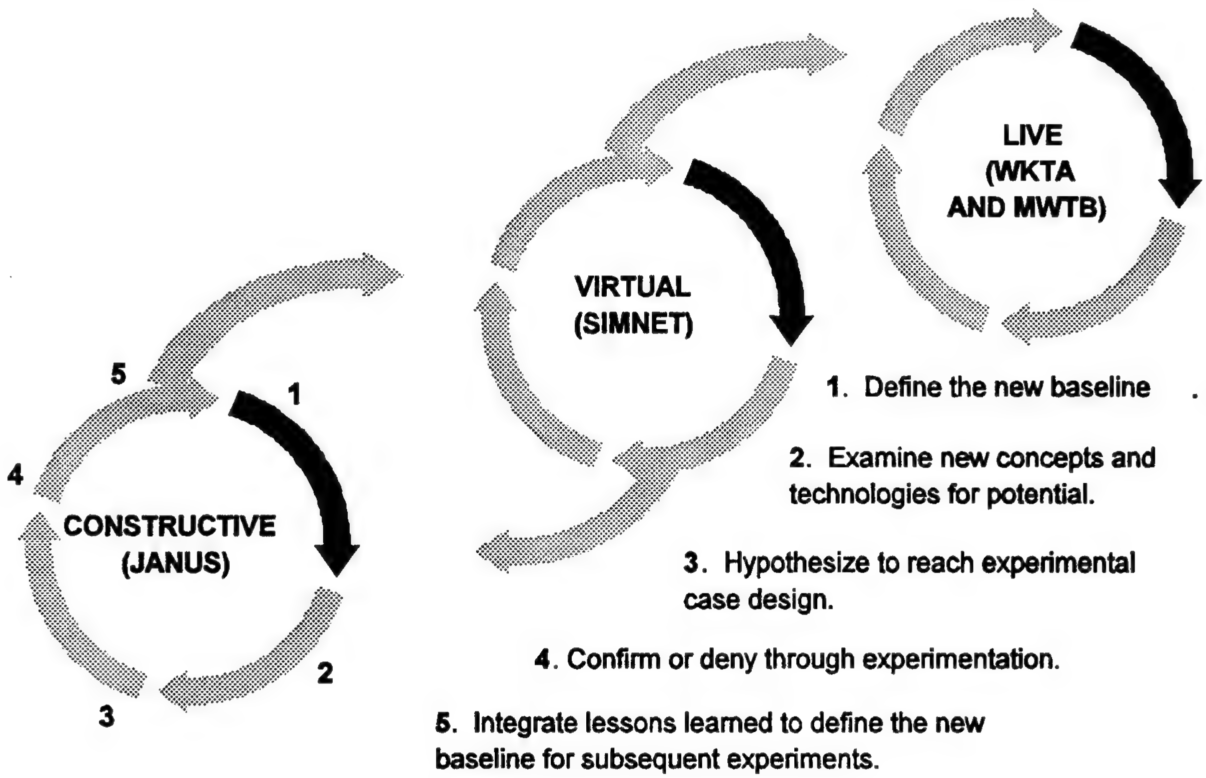

Figure 3. The spiral development process and a rolling baseline methodology as applied to Focused Dispatch (Adapted from U.S. Army Armor Center, 1995b).

The Method section of this report addresses how more responsive and formative research methods might be applied to the AWEs.

Notably, the Army is not alone in its efforts to change, to redesign itself. A recent hallmark of American business is the attempt by major organizations to manage their own radical redesign. Their experience suggests that such change is frequently uncertain, incomplete, and often demands action without complete information or in-depth analysis (Nadler, et al., 1995). Key lessons learned by these organizations bear on the Army's Force XXI efforts and underscore the need for responsive and pragmatic methods to evaluate and implement change: research methods that proactively engage and enable the ongoing process of change versus reactive and indifferent methods.

In sum, the military research community's recent concern with improving its methods and services reflect its awareness that the nature and context of military research issues have changed. Hollis' (1995) assessment is that the changing climate

in military research parallels the severe changes recently experienced in the national security environment: Things have become more "fuzzy." Hollis cites increasing demands for "near-real time analyses," quick-reaction studies, and less use of combat models. Future evaluations, he predicts, will require multidisciplinary teams and an increase in the use of Distributed Interactive Simulation (DIS). A prediction confirmed by the AWE's, including their application of DIS-based evaluations (Cosby 1995; Sikora & Coose, 1995).

Research Challenges in the AWEs

The AWEs typify the Army's emerging need for more pragmatic and responsive research methods to address the changing climate of military research. This brief review underscores two of the basic challenges AWE evaluators face in devising research methods for the AWEs: macro-level or "system-of-systems" complexity, and the developmental state of the focal AWE advanced information systems. These challenges reflect Joint Venture's formative objective and should decisively influence the types of research methods appropriate to the AWEs.

System-of-Systems Complexity. By design, the AWEs are global assessments of numerous new initiatives and their impact across the DTLOMS, with concurrent variation in nearly all variables (U.S. Department of the Army, 1995f). Difficulties in conducting research on military systems are due partly to the many variables present in any operational setting. "Even for single system tests, such as a single-seat fighter, the number of factors influencing operator performance is numbing" (Meister, 1987, p. 1294). More numbing is the fact that each AWE comprises numerous systems and subsystems, such as digital command and control devices in ground and air weapon systems, battle command vehicles, mobile command posts, and related combat service and service support elements.

For example, it was once estimated that the Task Force XXI AWE (see Figure 1) scheduled in March 1997 at the NTC would "involve 110 new initiatives, 53 new systems, and more than 20 new software applications" (Hewish, 1995, table). More recent estimates have pared the numbers cited by Hewish, but this brigade-level AWE retains a system-of-systems' complexity. Even conventional, non digital, brigade-level operations entail substantial numbers of weapon, service, support, and information systems controlled by soldier-in-the-loop participants. In

addition, Task Force XXI's digital focus may introduce an estimated 1,200 Applique components, this AWE's advanced information system centerpiece, and their varying level of connectivity with more than 40 different types of supporting digital systems (U.S. General Accounting Office, 1995). The problem space is further confounded by predicated interactions between many of these variables, including the synergy of horizontally and vertically integrated Battlefield Operating Systems and fully combined arms operations.

In sum, a basic research tenet is that substantiating a true difference is much more difficult than "finding" no difference. The potential for failure--negative, null or anemic results--is extremely high for large-scale evaluations (Boldovici & Bessemer, 1994; Johnson & Baker, 1974). A rolling baseline strategy, however, is predicated on the difference between current AWE performance and prior baseline performance. Notably, this strategy will not "benefit" from conclusions of no difference, faulty or not. While traditional and established programs are rarely evaluated, "it is the more venturesome program that bears the brunt..." of negative results (Weiss, 1972).

An important, if unstated, purpose of most research is to avoid failure. Reasons for such failures are legion and often traceable to unresolved method issues that are fundamental to all evaluation. Such fundamental issues include: vagueness of purpose, unrealistic claims, a multiplicity of objectives and issues, inadequate sample sizes, uncontrolled variation, presumed but unmet capabilities, inadequate training of participants and data collectors, and blunt or inappropriate measuring tools. Particularly, summative expectations that the AWEs will "prove" something must contend with their macro-level complexity. Resolution of such issues and closer adherence to fundamental evaluation methods are key to meeting summative and formative expectations.

Developmental Systems. The AWE's information technologies are in the early stage of development. Moreover, cost and a deliberate intent to exploit new and emerging technologies extend the developmental process for Force XXI systems. Past AWE evaluators candidly admit to sundry system limitations such as the availability, maintainability, and compatibility of the many different digital information systems employed (U.S. Army Armor Center, 1994, 1995a). AWE combatants and supporters have struggled to overcome such limitations and the inadequate and

complicated interface designs of their developmental prototypes. For example, AWE combatants and supporters have had to input repeatedly the same information due to system crashes or the need to manually transfer information between incompatible systems. Such limitations increase operator workloads and may decrease unit performance dramatically.

The information-based nature of Force XXI's information technologies dictates that soldiers and system operating procedures are critical factors in an AWE. The AWE's emphasis on information processing and management requirement puts soldiers and procedures directly in the operational loop, particularly for developmental systems. Invariably, training programs and operator manuals lag system development and are quickly obsolete when systems are revised or new systems are introduced. Only after users are well trained and practiced in the fundamentals of system operation can they progress to identifying, learning and applying higher-level TTPs for system employment. Additional documentation, training, and practice are required for AWE evaluation efforts directed, for example, at comparing variant TTPs, or TTP refinements designed to further leverage a potential system capability.

In sum, another research tenet is that the precision of resultant evaluation data is dictated by the stage of system development (Meister, 1965). The argument for addressing formative issues initially, emphasizes that precise and reliable data for summative conclusions, even model revisions, are not readily obtained with developmental systems. Notably, AWE developmental systems are not instrumented to provide more precise data automatically (U.S. General Accounting Office, 1995). The continuous and cross-unit nature of many AWE measurement issues, such as a common picture of the battlefield, are not accurately or easily captured by manual data collection procedures.

Developmental systems also raise concerns about obsolete and inappropriate measures. First, measures originally included in a rolling baseline may become obsolete as the AWEs progress and substantial changes occur in doctrine, organization, and training. Nevertheless, a driving concern of Joint Venture and the AWEs is to fully exploit the emerging, even unforseen, potential of advanced information systems. Second, actual versus presumed capabilities, such as the common picture available to

AWE participants, should guide AWE evaluation decisions about purpose, scope, and the avoidance of inappropriate measures (W. M. Parry, personal communication, 23 April 1996).

Experimental Conditions

The conditions of macro-level complexity and a reliance on developmental systems are integral to the AWEs. These research challenges constitute the prevailing climate, the experimental conditions, of the AWEs. Particularly, the conditions for field-based or live simulations by digitally-equipped battalion, brigade, and higher units in rotations at the National Training Center. These conditions are an AWE's "administrative" reality (Thompson and Rath, 1974) as determined by Army leadership and AWE directors, such as the Battle Labs. Realistically, at this level, the AWE's are army warfighting exercises in the spirit of the Louisiana Maneuvers (Merritt, 1995). Exercises regimented to spur the very hard and complex work of building the Army's Force XXI capabilities.

Experimentally, such conditions suggest AWEs are practical experiments in the service of discovery. AWEs are not fine-tune tinkering. They target the systemic, wholesale and concomitant changes essential to Force XXI. They are grounded by a practical and direct approach to discovery and development (Gregory, 1928).

The pragmatic need to conduct practical experiments was recognized by the primary architect for Joint Venture, General Sullivan: "I can tell you about a Grecian urn, but until you hold one in your hands, until you can really see it, even touch it, the magic is elusive. So it is with change. In some cases you have to create prototypes in order to create disciples." Sullivan also recognized the importance of labels assigned to the Army's Force XXI development activities. "First we called them demonstrations, and that didn't fly because we knew if it was a demonstration, that wouldn't be satisfactory for any of us. It would be an experiment" (Army, 1995).

Similarly, General Franks stated: "...we created Battle Laboratories in TRADOC (Training and Doctrine Command) as a means to experiment--discovery learning with real soldiers in tactically competitive environments--driven by the ideas of where battle is changing" (Army, 1994). In common usage, the word 'experiment' means to devise something for real world testing,

such as a policy or procedure, a trial balloon, a flying machine, or a digitally equipped force.

The Army's leadership appears to understand that the AWEs are about discovery learning and that they are formative exercises to see what works and what doesn't. Reportedly, a primary reason for changing "Advanced Warfighting Demonstrations" to "Advanced Warfighting Experiments" was to establish more realistic exercise conditions. Conditions that tasked real soldiers and equipment with the complexity and nascent capability of information-based operations.

Accordingly, the Chief of the Mounted Battlespace Battle Lab (MBBL) in charge of the Focused Dispatch AWE, iterated the Army perspective that this was a practical experiment. At the kickoff meeting for Focused Dispatch, he stated: "I don't expect everything to work. This is an experiment, not a demonstration. The fact that something does not work, whether equipment, training or TTP, may be just as important as the fact that some things do. We may learn just as much from those failures" (G. P. Ritter, personal communication, July 13, 1995). Before and during this AWE, MBBL leadership stressed that Focused Dispatch was not designed to prove the value of digitization. During the live-virtual exercise of Focused Dispatch, the MBBL Chief commented: "We are experimenting about experimenting" (G. P. Ritter, personal communication, August 17, 1995). The final Hot Wash slide on methodology for Focused Dispatch stated: "Reinforce <u>experiment</u> versus demonstration versus test" (U.S. Army Armor Center, 1995b).

In contrast to the open-ended nature of the AWEs and practical experiments, the scientific experiment is a precisely controlled research method (Cook & Campbell, 1979). Conditions for AWE scientific experiments are simple, but exact: random assignment of participants to an experimental and control group, and strict control over all factors extraneous to the causal variable of interest, such as the digital capabilities anticipated for Force XXI. Macro-level AWEs, however, afford neither random assignment nor strict control over experimental conditions. AWE participant assignments are nonrandom and typically made to only a single AWE experimental group. The AWEs' scope and complexity introduce myriad and dynamic extraneous factors, or variables, that contest reliable and valid outcomes, particularly for more summative force-level MOEs.

While the AWE's MEMV methodology acknowledges these research challenges, it may not adequately address them. The analytic plan for the MEMV method admits that the Army's warfighting experiments, to include AWEs, have inherent differences in design and structure, and a multitude of issues and objectives that change within and between AWEs. Rather than empirical control over conditional differences, or a systematic design to reduce variation, the plan proposes to document such differences (e.g., issues, design, structure, equipment) and conditions unique to each exercise, in a relational data base. The documentation of all exercise conditions is essential for interpretation of findings. However, it is no substitute for the controls required to establish commensurate conditions, systematic changes in treatment, and definitive or validated outcome improvements.

In sum, a more comprehensive research strategy might help meet Joint Venture's objectives to improve and prove force capability. This strategy might initially identify and apply research methods for the process of improving the force to better ensure the subsequent objective of force improvement is attained. Appropriate methods should responsively accommodate the AWE's inherent complexity and reliance on developmental systems. The research strategy for applying such methods might explicate how evaluation findings are implemented. Additionally, this strategy might address the more stringent methods and experimental conditions required to meet Force XXI's summative objective.

Program Evaluation: A Continuum of Research Methods

"Evaluation is an elastic word...," a word that stretches to cover methods and findings of many kinds (Weiss, 1972). In research, that elasticity affords access to a continuum of research methods as required to meet the purposes of the evaluation. Program evaluation research is traditionally characterized by its ability to press the tools of research into service to reach more accurate and objective findings and answers. It is particularly appropriate when the outcomes to be evaluated are complex, hard to observe, and made up of many elements reacting in diverse ways. For this report's emphasis on program evaluation methods, the word program equates to Joint Venture's program of research and development with the AWEs as the principal activities.

Program evaluation's continuum of methods includes formative and summative evaluations (Scriven, 1967). For this report,

formative evaluation methods are equated with the more responsive methods required for practical experiments designed to improve force capability. Summative evaluation methods are equated with the more exacting methods required for scientific experiments designed to prove force capability improvement.

The gap between practical versus scientific experiments is neither slight nor fully appreciated in the context of the AWEs. While both are vitally important forms of research, they possess differing purposes and, too often, differing personalities. As Mitroff (1985) comments on doing useful research:

> Very few academics would know how to recognize and handle a 'real world' problem if it bit them. Conversely, very few practitioners would know how to conduct systematic research... (p. 22).

Program evaluation's continuum of research methods provides an accepted and powerful strategy for bridging that gap. A program evaluation strategy could encompass the complementary and differential methods required to meet the AWE's formative and summative objectives.

Summative Evaluation Methods. Summative evaluations are conducted to generate information on the status of goal attainment, primarily for an external audience. Such evaluations are generally conducted after a program or system is fully developed and assess the effect of the subject system, such as digitally equipped forces, on outcome measures of performance. Reliable and valid research methods are needed to establish goal-related findings robust to an adjudicative audience wanting the AWEs to prove something. The AWE's macro-level complexity and reliance on developmental systems counter the exacting methods required for establishing robust findings based on force-level, soldier-in-the-loop MOEs. While this report focuses on AWE formative evaluation methods, a strategy for conducting summative evaluations in support of Force XXI is briefly addressed.

First, the AWEs do not preclude controlling their complex of conditions to validate key results. Appropriately robust methods almost invariably impose strict controls and afford systematic variation to precisely determine what works and what doesn't. For the AWEs, such control might be best obtained by dedicating some portion of an AWE to subexperiments. True subexperiments require careful application of the methods for a scientific

experiment devised to control conditions, however, not mere miniaturizations that maintain an AWE's macro complexity. Methods for such summative purposes are referenced repeatedly in this report (Boldovici and Bessemer, 1994; Dewar et al., 1994; Leibrecht et al., 1994), and throughout the Method section, particularly under Structured Scenarios.

Second, controlled experiments designed to prove something about the effect of advanced information systems on force capability are not confined to the AWEs. Joint Venture's Campaign Plan and this report stress a too forgotten point: the AWEs are not stand-alone evaluations (see Figure 1). Force XXI objectives should be supported by related Army RDA&Tng programs. For summative objectives, these RDA&Tng programs and the Army's Battle Labs could provide controlled settings for conducting scientific experiments. These ongoing experiments could contribute directly to the body of evidence supporting the AWEs and Joint Venture's fundamental hypothesis.

Formative Evaluation Methods. Formative evaluations are conducted to generate information on the process of goal attainment, primarily for an internal audience. Such evaluations are generally conducted throughout the developmental stage to help form or improve the system for those who use it. The focus of formative evaluations is on intermediate goals, and the purpose of formative evaluation is to play a productive role in pursuit of those goals. Responsive and informative research methods are required to understand and improve the system, to obtain findings useful to system developers, designers and operators. Perhaps, the most useful purpose of evaluation is to identify aspects of a system or program where revision is desirable (Cronbach, 1964).

Intermediate and process measures, even when qualitative and subjective, may provide a more useful basis than quantitative outcome measures for improving force capability. Process and intermediate measures greatly increase the probability of obtaining reliable and valid measures (Dewar et al., 1994; Sackett & Larson, 1990), and could do the same for AWEs. Compared to the numerous data points collectable on intermediate measures such as messages transmitted or targets detected, MOEs such as force-exchange ratios may result in only one data point per AWE exercise.

Answers to many of the key issues confronting Joint Venture require explanatory and diagnostic information rarely provided by "raw numbers" compiled for MOE comparisons. An AWE focus on process and intermediate measures should identify many important and correctable deficiencies, by source and type. This level of specificity should result in useful lessons learned. Useful in that they precisely diagnosis a problem and provide clear and accountable guidance on lesson implementation. Implementation should directly remove or overcome the identified deficiencies or shortcomings obstructing the capabilities anticipated for Force XXI. Initially, such AWE process and intermediate measures are needed to achieve force capability improvement. Ultimately, they are needed to substantiate conclusive MOE results of such improvement and achieve further improvement.

Fundamental Evaluation Methods. All evaluation activities are essentially similar, and fundamental research methods that relate data collection to evaluation goals apply to summative and formative efforts. A basic concern is that the AWE methodology address the fundamental evaluation methods that underlie reliable and valid measurement. Such fundamental methods establish AWE evaluation preconditions and conditions essential to meaningful and useful findings.

Fundamental AWE method issues include a multidisciplinary evaluation team, the purpose and scope of the evaluation, and precise data collection methods. AWE method fundamentals also include structured exercises, functional tests, and trained participants and data collectors. These fundamental AWE method issues and method recommendations for addressing each issue are considered in the Method section.

In sum, the quality or "goodness" of evaluation data depends on successful accommodation to the situation. This Introduction section stressed why evaluators should adapt research methods that accommodate contemporary military research, particularly the AWEs. Research quality in a more general context was addressed by Thompson and Rath (1974):

> ...'good' research is that in which the researcher bases his choice of method on the degree of his initial uncertainty and is careful to disclose the accompanying degree of uncertainly in his results. p. 243

Method

The AWE research methods proposed in this section correspond to the assumption that the AWEs are primarily designed to improve force capability through the integration of information-based technologies. This assumption underscores that the AWEs are a series of practical experiments, advanced warfighting exercises, regimented to spur the difficult and complex work integral to Force XXI. The primary AWE research challenge is a successful adaptation to this climate that provides useful solutions to the formation of the future force.

This section addresses fundamental and formative evaluation methods for the AWEs. It identifies a set of twelve AWE research method issues and provides corresponding method recommendations to address each of the issues raised, as summarized in Table 1. Fundamental versus formative issues and methods are separately identified in Table 1. Their order of presentation in this table reflects a logical sequence for addressing each issue and its related recommendations.

The AWE method issues addressed in this section are based, in particular, on the author's observations of two AWEs, Desert Hammer VI and Focused Dispatch. In general, these issues are based on the work and observations of other researchers concerned with evaluation methods, as referenced in this report. Notably, these recommendations do not constitute a complete methodology for conducting AWE formative evaluations. The method issues raised, however, identify twelve key building blocks essential to that methodology. In support of each issue, the recommendations direct AWE evaluators to the guidance and tools provided by other evaluators that might assist the Army's effort to form the future force.

Eight of the twelve issues addressed in this section are fundamental issues, common to all evaluation. All evaluation activities are essentially similar, that activity is the collection and combination of performance data in relation to evaluation goals (Scriven, 1967). Adherence to such fundamental methods is essential to obtain reliable and useful data in relation to summative or formative goals.

The four remaining method issues considered in this Method section are more unique to AWE formative objectives. The method

Table 1. AWE Method Issues and Recommendations

Issue	Recommendation
Multidisciplinary Team	Build multidisciplinary evaluation team: research, operational, system, technical and training skills.
Purpose of Evaluation	Align evaluation team with the primary purpose of the AWEs (e.g., to improve versus prove force capability).
Scope of Issues	Guide administrators' determination of a limited, realistic set of key issues and deliverables (e.g., products).
Evaluation Methods	Define research methods fitting to the evaluation's primary scope and purpose.
Functional Analysis*	Develop detailed understanding of equipment, personnel and procedures.
Task Analysis*	Analyze key tasks, conditions and standards for system operation.
Process Measures*	Develop measures that assess the performance required to achieve outcomes, versus outcomes per se.
Performance Model*	Develop models of performance for key tasks (e.g., for advanced information technologies).
Structured Scenarios	Structure scenarios and exercises with conditions that predicate repeated performance of key tasks.
Functional Tests	Conduct "loaded" tests of system functions and simulation utilities.
Trained Participants	Train participants to proficiency in performance of key tasks.
Trained Data Collectors	Ensure data collectors understand the focal system and requirements for task performance and data collection.

* Indicates method issues and recommendations more unique to formative evaluation.

recommendations for these issues are especially appropriate for attaining some of the key intermediate objectives of Force XXI. Method issues and recommendations more uniquely related to formative evaluations (Bloom, Hastings & Madaus, 1971; Fitz-Gibbon & Morris, 1987), such as the focus on process measures, will receive special emphasis in this section.

In preface, the power of formative evaluation derives from its contribution to *understanding* the system or program being developed (Scriven, 1967). The formative evaluation's focus on intermediate goals, therefore, necessitates a more detailed and direct examination of key system or program factors including: the optimal allocation of system functions and operator workload, actual versus espoused system capabilities, participants' understanding of the system, the content and structure of participants' training to operate the system, and the TTPs of system employment. To exemplify the detailed understanding required, the method recommendations provided focus on TTPs for exploiting information technologies, and on the process of maintaining a relevant, common picture of the battlefield for Force XXI combatants and supporters.

Information gleaned from direct examination of formative factors, may: identify the precise location of shortcomings in the system, distinguish between importantly different explanations of success or failure, suggest probable causes for such failures, indicate a lack of practice in basic or system-specific skills, and inform improvement (Scriven, 1967). In contrast, aggregate or summary outcome scores provide little toward our understanding of the system or how it might be improved. "It is the *nature* of the mistakes that is important" in evaluating a system or program, and in revising it (Scriven, 1967).

A key concern is how findings from the formative evaluation portions of an AWE might be implemented to support Force XXI objectives. This concern is directly addressed in a subsequent section, Utilization of Findings, that suggests "living products" should implement lessons learned. Method recommendations made throughout this Method section, therefore, embed a mechanism of expanded AWE evaluation teams for maintaining and sustaining these living products across the AWEs and Army-wide Force XXI efforts.

Multidisciplinary Team

The broad scope and complexity of the AWEs underscore Hollis' (1995) call for multidisciplinary research teams with technical and practical expertise across systems, operations, plans and evaluation requirements. In essence, an AWE evaluation team should include representatives from all key areas of the current AWE and from Army-wide Force XXI efforts related to that AWE. The Department of Defense and the Army's recent efforts provide a useful model for forming and applying multidisciplinary teams that integrates versus fragments developmental efforts (Langford, 1995). One characteristic of these multidisciplinary teams, such as an Integrated Product and Process Team (IPPT) for streamlining the acquisition process, is a continuity of involvement throughout a system's life cycle, from concept to fielding. A second characteristic fundamental to the team concept, and IPPTs in particular, is coordinated and collaborative effort.

This report stresses that such continuity within and between AWEs should be a key concern in forming an AWE evaluation team. The AWEs should not be stand-alone evaluations. Such continuity is essential to build the expertise needed for development and refinement of the advanced information technologies that are the foundation of Force XXI objectives. The sustained sharing of expertise provided by each team member is also essential to the team's overall need to develop a broad and detailed common-knowledge base. This knowledge base is needed to understand and leverage the system-of-systems' potential for synergistic improvements in fully combined arms operations. To achieve this continuity of effort for AWEs, this Method section recommends and illustrates how an expanded team of experts from related Army-wide programs and agencies might supplement each AWE's core evaluation team.

Team coordination and collaboration is facilitated by holistic and interdependent goals directly linked to explicit products. A formative/developmental focus on the process of achieving Force XXI objectives provides an excellent basis for interdependent goals and integral products that knit the evaluation team directly into the AWEs. In contrast, summative evaluation often alienates evaluators from the soldiers, equipment and procedures they are evaluating. For example, consider the roles of "independent" and external evaluators. While such roles may reduce the bias frequently aroused by

adjudicative assessment (Webb, Campbell, Schwartz, & Sechrest, 1966), they discourage the collaboration required for formative assessment and coordinated products.

Guided by these desired team characteristics, the actual formation of an AWE multidisciplinary evaluation team might begin by considering team leadership. An evaluation leadership that directly supports and answers to the actual AWE directors, such as the Battle Labs. Based on the assumption that soldiers and procedures are vital factors in an AWE, the leadership of the AWE evaluation team requires human performance specialists with demonstrated expertise in the evaluation of human behavior in complex systems. The information-based nature of digital systems, particularly the AWE's developmental and manually intensive systems, promotes soldiers and operating procedures to the forefront. Formative research methods should detect and direct needed improvements in operator training, procedures, performance, and workload. The need for human performance specialists is also stressed for an AWE's empirical evaluations based on soldier-in-the-loop simulation, live and virtual. Portions of an AWE based on constructive simulations, on the other hand, might be guided by analytic specialists focusing on organizational issues, for example.

Leadership guidelines for evaluation teams urge that team leaders must be empowered with the authority to make decisions within the evaluation team context (Plott, LaVine, Smart and Williams, 1992). Within the AWE's context, these leaders must work in concert with AWE directors such as the Army's Battle Labs to initially resolve fundamental issues, such as the purpose and scope of the evaluation, that will guide development of evaluative methods.

Next, the general roles and responsibilities of the core and expanded evaluation teams should establish a framework for coordinating formative evaluation and implementation efforts within and between the AWEs. Particularly for formative evaluation, the core evaluation team should not be limited to "evaluators" per se. The more traditional core AWE evaluation team should be buttressed by expanded teams of advisors, specialists, and assistants (e.g., system experts, task analysts, and trainers). The core team and its activities may exclusively serve a subject AWE, but members and activities of the expanded teams should provide recurrent service across the AWEs. Members of the various expanded teams should be drawn from related Army

RDA&Tng programs and agencies. These expanded team members should provide DTLOMS-wide representation and expertise related to the AWE's focal issues and advanced information systems. Resourcing arrangements similar to the ATD's, ACTD's, and IPPT's will be needed for these expanded team members' extended and recurrent involvement in the AWEs (U.S. Department of the Army, 1996b).

The primary product of the core team should be an Evaluation Support Package (ESP) that precisely specifies evaluation activities and procedures for the AWE. For the team's formative goals, this ESP should stress intermediate and process measures that are explicitly defined and checked against data collection resources. In addition to their expert contributions to the core team's knowledge base, the members of the expanded teams should provide and tailor explicit products from their respective programs into the AWE. These products should include: doctrinal and organizational literature, system specifications, functional analyses, performance models, TTP manuals, and Training Support Packages (TSPs) (U.S. Department of the Army, 1996c). In turn, lessons learned during the AWEs should be implemented back into the expanded teams' respective products, and these products exported back to their agencies and programs.

The cohesion of the core team may be facilitated by the strategic sharing and learning of a common and specialized knowledge base that is needed for understanding AWE system and evaluation factors. To build this knowledge base, every effort should be made to ensure the expertise of the team members includes experience of direct relevance to the systems and methods to be employed. A reliance on notional expertise, or background in unrelated areas, may only undermine establishment of the team's knowledge base and impair evaluative efforts. Given the expertise required, directed collaboration on explicit and integrated products should foster cohesion within the evaluation team and between this team and AWE participants, sponsors, directors, and supporting players.

Critical members of the core team include representatives of the AWE's directors, such as a Battle Lab, who would maintain coordination between the evaluation team and the AWE directors. While communicating the goals and objectives of the directors and sponsors to the evaluators, these representatives should also iteratively inform the evaluators about scheduling, resources and constraints, planned and revised. These representatives must

also liaison back to the directors the evaluation team's recommendations, such as a realistic scope of issues and acceptable methods of assessment, given AWE resources and constraints.

Other key team members include technical and operational experts for each system included in the AWE. Such experts should provide the detailed knowledge of system functions and operator procedures. This delineated understanding of an AWE's advanced information technologies is essential for formative evaluation and improvement. These technical experts should provide and maintain the team's accurate understanding of the actual versus projected capabilities and limitations of each system. Operational experts should provide and/or review task-based analyses of the exact techniques and procedures required for system operation in the context of realistic military scenarios. The core evaluation team should be augmented by expanded teams of technical and operational experts who provide and revise important products, such as the functional and task analyses described in subsequent sections.

Although some technical and operational expertise is generally available during the AWEs, it has not been routinely aligned to directly support the evaluation team during and between AWEs. In fact, the developmental nature of AWE component systems and their methods for employment (e.g., TTPs), have greatly limited the expertise available from civilian and military team members. Notably, the Army is developing a data base that identifies digital experts who could be referenced in the selection of future evaluation team members, such as system and operational experts. The challenge of growing and sustaining a team of such experts, available to both the AWEs and related warfighting experiments, requires that such membership does not adversely affect their career progression, particularly for military members.

Other key members of the evaluation team include technical and site representatives from the live, constructive and virtual simulation settings to be used during the evaluation. The expertise of these members in identifying the advantages and limitations of simulation unique to their setting is needed to assess the feasibility of evaluation objectives, to identify key task conditions and system capabilities that can be simulated, and to inform the team about measurement resources and issues within their simulation environments. Such experts can also

assist substantially in the development of training and evaluation schedules within their settings.

Additional key members of the core evaluation team should be representatives from the training team and the data collection team. Trainers and evaluators must concur on training and evaluation objectives to ensure that participants are proficiently trained on the precise aspects of performance required during AWE exercises. Trainers and evaluators should collaborate with appropriate team members to achieve a clear and detailed understanding of key system functions and operations, and to assure that scenarios used during training and evaluation effectively and efficiently simulate the task conditions associated with key task performance. The AWE's reliance on novel equipment and TTPs, underlines the need to ensure that participants receive a structured training program that addresses and exercises all key tasks slated for evaluation. Finally, the core team should from the onset include representatives from the expanded data collection team. The critical role of trained data collectors is the last issue considered in this Method section.

The AWE emphasis on live and realistic exercises conditions underscores the requirement that the evaluation team include members from the Combat Training Centers (CTCs). The essential role of the CTCs in Army experimentation is articulated by Webster (1995) who stresses that the CTCs should be involved early in AWE planning and that adequate training of CTC trainers, supporters, and data collectors is key to an AWE's success. Despite the CTC's impressive array of instrumentation, Webster urges additional resourcing or compatibility with current CTC instrumentation to monitor, collect, and analyze new equipment data. Overall, CTC representation on the evaluation team would help ensure that AWE evaluation methods are meshed with CTC personnel, equipment, and the Rules of Engagement.

In sum, the core evaluation team bears ultimate responsibility for establishing and implementing an AWE's evaluation methods. The predetermined nature of the AWEs, particularly their macro-level complexity, underscores that the evaluators' methods, fundamental and formative, must support the constraints and purposes of the Battle Labs and Joint Venture. The expanded team's concept that enables the AWEs to work in concert with and directly benefit from related Army efforts, firmly gears the team's evaluation efforts toward living products responsibly developed and shared Army-wide.

Purpose of Evaluation

The relationship of research purpose and method might seem evident, but it is not (Meister, 1987). Far too often evaluation efforts fail to specify what they are actually measuring and why. The reasons for such failures may originate in uncertainty about the purpose of the evaluation. Frequently, evaluations are couched in abstract purposes and gross constructs such as operability and reliability. For the AWEs, pervading constructs are lethality, survivability and tempo. The research community has not resolved the underlying component measures for such constructs or their realistic applicability to the AWEs, as designed and resourced. Evaluations inevitably flounder when evaluators do not press from the onset for very specific answers to "What is it we are supposed to measure?" (Meister, 1965).

The iterative and radical changes anticipated across the AWEs suggest that for each AWE evaluators may need to review and revise the measurement focus. An earlier AWE, such as Desert Hammer VI, may redirect a subsequent AWE, such as Focused Dispatch, to more closely assess training and TTPs. AWE evaluators may need to refine their measures during and between AWEs to ensure they are not measuring old TTPs for new systems. The competing demands for wide-ranging assessments in the AWEs macro-level context also urge AWE evaluators to clearly specify what it is they are <u>not</u> measuring.

The AWEs' purpose as a formative evaluation differs substantially from that of traditional scientific research. Adapting a comparison made by Weiss (1972), evaluation research is: pragmatic--used for decision making, rather than the mere accumulation of knowledge; programmatic--driven by the program goals, not the researcher's; partial--with respect to program goals versus "objective" or unbiased; progressive--in that improving and implementing the ongoing program are of higher priority than the costly and time-consuming methods frequently required for gaining dated defense of the program; and, personal --characterized by role conflicts that make program leaders and participants suspicious of spurious evaluative methods that may fail to support the program to which they are committed.

In sum, an emphasis on methods to iteratively improve force capability should guide the efforts of AWE directors and evaluators in defining the primary purpose of an AWE. This report's recommendations on fundamental and formative research

methods might foster congruence on how to pursue that purpose. Directors' and evaluators' concurrence on evaluation purpose is essential to achieving AWE evaluation objectives. The Army's Battle Labs and senior leaders ultimately determine the purpose of the AWEs. The evaluation team, however, should actively assist these directors in determining evaluation purposes that can be legitimately addressed given the resources and constraints of the AWEs. Once concurrence of purpose is thoughtfully achieved and thoroughly disseminated, the members of the evaluation team should fully align with that purpose.

<u>Scope of Issues</u>

The need for specific and feasible goals in complex, modern endeavors is recognized by the military's first principle of operations: "Direct every military operation toward a clearly defined, decisive and attainable objective" (U.S. Department of the Army, 1994a). The application of this principle to the formal requirements for AWE evaluation begins by parsing the general purpose of the evaluation into a discrete and attainable set of research objectives and related issues. Evaluators, like commanders, with unclear mandates should take the initiative to refine and define their mandate for consideration by higher authority, the AWE directors.

Illustrating the need for closure, a draft list of Joint Venture issues contains 189 separate research issues (U.S. Department of the Army, 1995a). Even this extensive issue list entails very broad and abstract research issues such as "What is the process to maintain a near real time relevant common picture?" Notably, this is a process issue, but answering such a global issue requires detailed specification of key system components, the measures needed, and how their data can be obtained. Outcome MOEs do not address such an issue. Explicit measurement definitions and resources appropriate to such AWE issues have not been adequately identified or developed.

A multiplicity of AWE research issues only blunts efforts to develop the specification required for more precise methods and measures. One strategy for reducing the number of potential research issues from a wish list to a subset of attainable objectives is to systemically review the evaluative requirements underlying each issue. By adequately formulating potential research issues or questions, researchers and directors would pass through a series of specifications that lend needed closure

to the identification of legitimate evaluation issues. Key steps in this formulation of research issues include: specifying the participants, setting, and variables of interest; explicitly stating what type of variable relationship (e.g., correlational, causal) is of concern; and, precisely defining the unit of analysis such as the individual, the crew, the unit, or the entire force (Sackett & Larson, 1990). A realistic culling of issues might also include specifying, at least draft-level, explicit definitions of measurement procedures, checked against data collection resources.

Classic strategies for reducing complexity of purpose include "shortening the chain" of variation, subexperiments, and controlling variability (Mann, 1972). A focus on AWE process measures may provide an effective method for evaluators to shorten the chain of variation between controlled and uncontrolled variables, and what is being measured during the AWEs. Intermediate measures of the moment-to-moment processes underlying TTP performance, for example, should provide information more directly related to TTP improvement than remote outcome measures. The use of AWE subexperiments to better isolate key variables from macro system assessment is strongly urged, and reinforced here by focusing on selected issues such as the common picture of the battlefield and the TTPs for employing information technologies. Efforts to reduce the variability in AWE evaluations should be guided by the work of other evaluators (e.g., Boldovici & Bessemer, 1994; Boldovici, 1995), and focused on the identification of method-based issues for directors' review. Additional methods for controlling AWE variation, such as trained participants, data collectors and structured scenarios, are reviewed in subsequent method recommendations.

Tactics, Techniques and Procedures (TTP). A key research issue that should permeate all AWE efforts is the set of TTPs needed to effectively employ digital technologies on the future battlefield (Rigby, 1995). The Battle Lab Experimentation Plan (BLEP) for Focused Dispatch, for example, stated that a primary purpose of that AWE was to refine TTPs for the Mounted Task Force (TF) of the future, TF XXI (U.S. Army Armor Center, 1995a). That statement identified the TF as the unit of analysis and could have provided a solid basis for narrowing the scope of this AWE to examination of these TTPs. In addition, that BLEP attempted to devise a set of TTP subexperiments for its live, virtual and constructive settings to address important process issues such as

"Call For Fire" routing alternatives and dispersed movement techniques.

Earlier versions of the Focused Dispatch BLEP attempted to further define the scope of effort by identifying a subset of key TTPs called "alpha" cases from the overall set of TTPs recently developed for a digital TF (U.S. Army Armor School, 1995b). For a subset of these alpha cases, variant or excursion TTPs labeled "beta" cases were also broadly defined. The BLEP originally proposed that each beta case variant would be compared with its respective alpha case. For example, the alpha case for Call For Fire entailed traditional routing options for digital requests for indirect artillery fire. The beta case variants for this TTP progressively simplified the routing of these requests, culminating with a direct "sensor-to-shooter" digital link.

In sum, the proposed alpha and beta case TTPs for Focused Dispatch provide an excellent example of how the scope of issues for AWE assessment might be honed to identify attainable research issues and objectives. Due to shortcomings in equipment, training, and documentation (Elliott, Sanders & Quinkert, 1996), these alpha-beta comparisons were not thoroughly tested during Focused Dispatch. Nevertheless, this AWE's focus on key TTPs is reinforced in these AWE method recommendations.

Evaluation Methods

Defining evaluation methods entails specifying "the procedures and activities used to collect and analyze a set of empirical data bearing on some question of interest" (Sackett & Larson, 1990). As an overview, the twelve method issues and corresponding recommendations in this section should contribute to the definition of AWE evaluation methods. For example, prior AWE method recommendations on forming a multidisciplinary evaluation team, defining the purpose of the evaluation, and narrowing the scope of evaluation addressed some of the fundamental issues that should be resolved before explicit data collection procedures and activities can be defined.

A focus on formative evaluation methods, directs evaluators to identify methods and measures required for a detailed understanding of the equipment, personnel and procedures needed to improve employment of future force technologies. If we apply Meister's (1965) definition of a complete system to the AWEs, the component parts begin to unfold: equipment, supporting

architectures for simulation settings and communications (digital and voice), participants, support personnel, procedures, training, DTLOMS, logistics, and technical data. As an example of the detailed knowledge base required to define evaluation methods for improving force capability, reconsider "What is the process to maintain a near-real-time and relevant common picture of the battlefield?"

Common Picture of the Battlefield. A hallmark capability anticipated from Force XXI information technologies is the ability to provide and maintain a "common picture" of the battlefield for combatants and supporters. Identifying and improving the processes underlying this capability is, therefore, an important issue for the AWEs and the Army's supporting research and development efforts. The common picture issue also illustrates how AWE evaluators can decompose decision-makers' abstract and subjective issues into concrete and objective data elements. Common picture elements, for example, provide objective grist for forming and assessing more elusive constructs, such as the mental picture or situational awareness of a commander.

Identifying and developing evaluation methods for this common-picture issue, as with most AWE formative issues, requires a multidiscipline knowledge-base about supporting systems, personnel, and procedures. General method recommendations for establishing and maintaining such a knowledge base are presented in subsequent sections on Functional Analysis, Task Analysis and Performance Models. Here, particular examples of the detailed knowledge required to address the common-picture issue are considered. To aid this consideration, Figure 4 depicts the key equipment and information links for maintaining a common picture during the live/virtual exercises of Focused Dispatch.

The first task in determining evaluation methods for the common-picture issue might be specifying the data and informational elements required for a generic battlefield picture. Component elements probably should include, at a minimum, traditional factors such as Mission, Enemy, Terrain, Troops, and Time (METT-T). Next, method determination might identify the different data sources for each element, the informational sources required for assimilating and interpreting that data, and the designated recipients of the common picture.

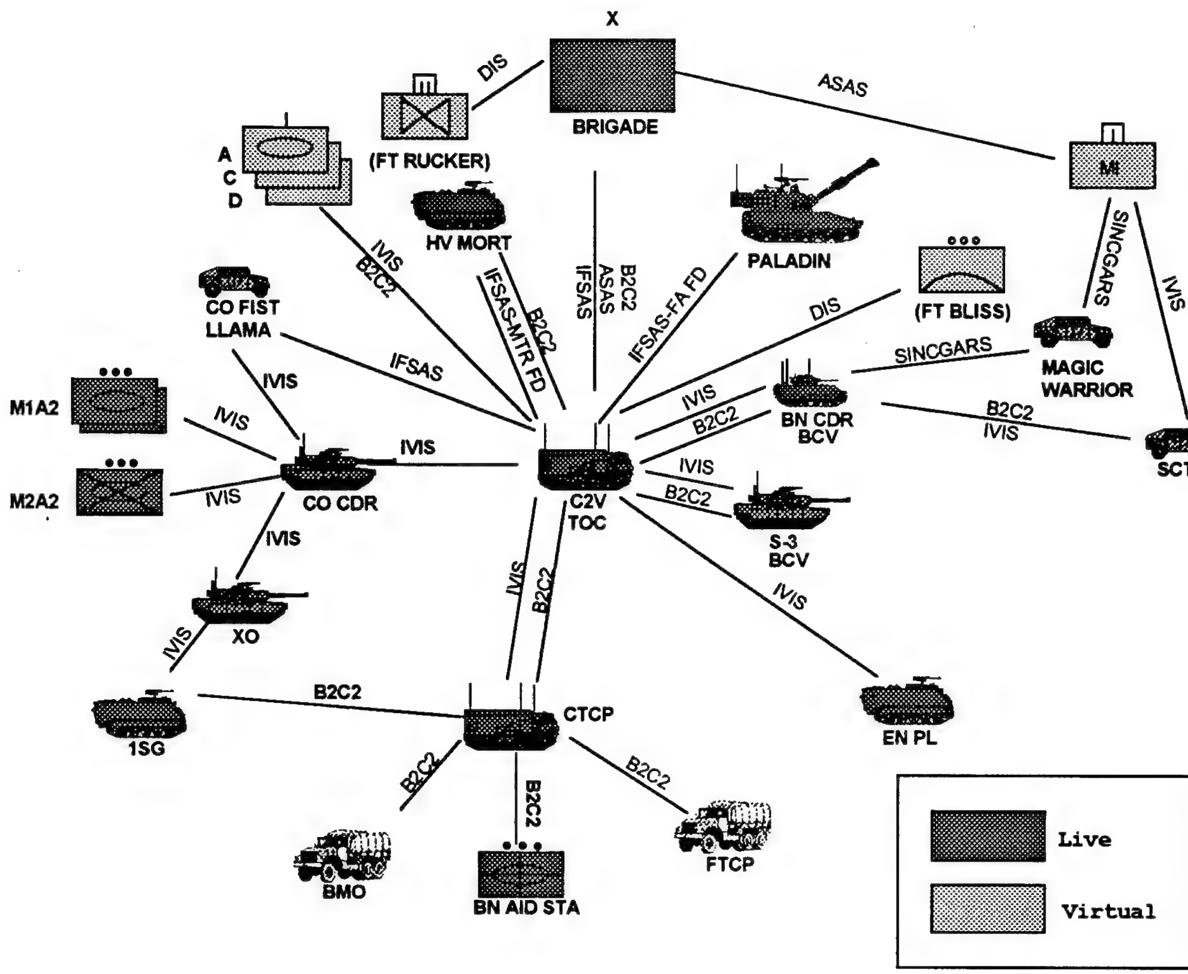

Figure 4. Equipment and communication linkages used in Focused Dispatch's live/virtual exercises (Adapted from U.S. Army Armor Center, 1995b).

Next, supporting functions and capabilities of the AWE's different informational systems used to provide the common picture should be identified to include: data formats for each element, message routing and distribution networks, transmission speed and accuracy, intersystem connectivity and compatibility. Finally, manual and automated tasks and procedures for acquiring, processing, distributing, storing, and updating each element of the common picture should be delineated.

Before specifying the data collection procedures and activities for this common-picture issue, precise operational definitions of supporting measures should be developed and checked against data collection resources. More general method recommendations on measurement and collection issues are addressed in subsequent sections, Process Measures and Trained

Data Collectors. Here, particular concerns with respect to the common-picture issue are considered. Explicit measurement definitions for this issue build from the previously defined information, such as informational elements, data sources, recipients, and procedures.

Next, the constructs of relevance and real-time might be clearly defined. Initially, considerations of relevance may require that different pictures are defined based on the informational requirements unique to each recipient's duty position (e.g., Figure 4). Eventually, user-based relevance requirements may entail definitions uniquely tailored to each recipient. Similarly, system-based accuracy and transmission speed as well as manual and automated processing and distribution procedures provide a realistic basis for quantifying and assessing "real-time" predications.

Once the above specifications and definitions are provided, attempts to specify the data collection elements for the common-picture issue are relatively straightforward. Determining the procedures for collecting the necessary data, however, requires careful consideration of available and appropriate manual and/or automated data collection instruments. Manual efforts to monitor and record the content and timing of each participant's common-picture updates during dynamic AWE scenarios are highly taxing and prone to error.

After data collection instruments and procedures for each issue are determined, evaluators should identify additional evaluation methods that will help assure meaningful and reliable data are obtained. Subsequent recommendations on structured scenarios, functional tests, and trained participants address many of the general requirements associated with collecting meaningful and reliable data. With respect to the common-picture issue, scenarios might be structured to ensure participants are required to repeatedly receive, process and distribute data related to each of the informational elements that constitute their relevant depictions of the battlefield situation.

Structured scenarios and exercises can precisely predefine an optimal picture that should be available to a recipient at any moment during a scripted battle. Optimal picture examples can then be directly compared with actual "snapshots" of the picture, as depicted on an operator's information display at selected times. More generally, structured scenarios greatly assist data

collector efforts to monitor, record and assess the processes observed. Functional tests help document the actual system capabilities for providing and maintaining the common picture, and identify shortcomings that impact the data obtained or processing anticipated. The training of participants attempts to ensure they are proficient in the multiple procedures and techniques required for generating, receiving and distributing common-picture data. Evaluations of their training should aid interpretation of the process data obtained, and provide useful information for improving training directed at the common-picture process.

In sum, meaningful improvements in system or force capability are generally based on a detailed understanding of underlying capabilities and limitations. If advanced information technologies are the "black box" basis for Force XXI improvement, evaluators must peer into that box by devising evaluation methods that more precisely disclose these capabilities and the processes required to achieve the capabilities anticipated. As indicated by the common-picture issue example, AWE formative evaluation methods should focus on understanding and improving these processes. Formative evaluation methods directed at key intermediate products could provide useful and essential information for achieving Force XXI capabilities.

Functional Analysis

Perhaps, the most critical aspect of common knowledge among members of the AWE formative evaluation team is their detailed understanding of system functions and how those functions are allocated between soldiers and machines. The system is "everything" required to perform the specified operation, Meister (1965) advises. Despite the imposing nature of this requirement, understanding the total military "system" established with the insertion of information technologies into combat, combat service and combat service support systems is essential to its improvement.

The primary system elements of equipment, personnel, and procedures are key to that understanding. The interaction of these elements is best established by a functional analysis of the complete system. An adequate functional analysis allocates these functions, describes the tasks performed by personnel and equipment, and specifies the criteria used in system development

(Meister, 1965). This analysis should address intermediate goals and criteria also, in the case of formative evaluation.

Again, the common picture issue exemplifies the formative role functional analysis might play in improving and evaluating an AWE's advanced information systems. A functional analysis directed at the common picture issue might begin by identifying for each of the METT-T categories what informational elements are functionally supported by each of the subject digital systems employed in an AWE. This initial analysis of equipment should detail the nature of that support to include: data formats for each element, message routing and distribution networks, transmission speed and accuracy, and intersystem connectivity and compatibility.

Next, the functional analysis might precisely identify how the functions and tasks for depicting and updating that picture on each type of display are allocated between soldiers and equipment. These analyses should identify functional voids and shortcomings in the common picture process. Useful findings might disclose, for example, unacceptable tradeoffs (e.g., time-consuming and repetitive manual procedures) that degrade the accuracy, clarity, completeness and timeliness of the common picture, as achieved versus anticipated.

Notably, the functional analyses just described can and should be conducted <u>before</u> an AWE. This analysis is an example of a living product provided to the core evaluation team by the expanded functional analysis team of system developers and operators. This product could guide an AWE's evaluation efforts directed at improving the common picture process, for example. Findings from the AWE might include a useful data base of lessons learned that precisely identify correctable shortcomings in this process and direct post AWE implementation of lessons learned into refined products, including: functional and task analyses, system requirements, and training packages.

One source of authoritative specification for an AWE's system and components is generally provided by the BLEP (e.g., U.S. Army Armor Center, 1995a). The Focused Dispatch BLEP provided brief descriptions of the overall system, particularly equipment and personnel, required for conduct of the AWE (U.S. Army Armor Center, 1995a). It also provided brief descriptions of each experiment, the experimental setting(s)--live, virtual or constructive--and their supporting architectures. System

limitations and overall complexity, however, may force repeated modifications in a BLEP before and during an AWE. Daily AWE modifications might include changes in mission, tasks and TTPs; changes in issues based on a prioritized schedule; and, revisions in architectures to support the currently scheduled issues and tasks, or to overcome unforeseen limitations.

In sum, a thorough functional analysis of an AWE-level system is a difficult but potentially rewarding method for evaluation. The core evaluation team will need the assistance of an expanded functional analysis team. This expanded team's involvement in information-based technologies and related Army research and development efforts may include previously conducted functional analyses for AWE subsystems and prior AWEs that should be adaptable to the current AWE. Their knowledge of function and task allocations between equipment and personnel should directly support, for example, evaluation and refinement of the TTPs under investigation. In particular, the evaluation team's knowledge of differences in functional allocations between conventional and digital system variants should assist in the identification of important AWE issues for improving capability. Overall, this expanded team should contribute to the AWE and provide a direct conduit for implementing AWE lessons learned directly into related research and development efforts.

Task Analysis

Traditionally, the task analysis requirement is regarded as the final stage of functional analysis. In support of an AWE formative assessment, task analysis is included as a separate methodological recommendation to emphasize the importance of a detailed understanding of task performance to the improvement of force capability. The identification of soldier-machine operations required to perform system functions provides the basis for establishing the task requirements for each system. The review of mission requirements by mission segment, moves the analysis toward a detailed examination of the specific operator tasks and the input-output relationships for task performance and evaluation. Task descriptions serve as the bedrock for determining the operator behaviors--perceive, discriminate, decide, and manipulate--to get from the input to the operator's output as well as the required system output (Meister, 1965).

As with the overarching functional analysis for a digital system, a detailed task analysis to support AWE evaluations may

not be available to the evaluation team. For example, supporting documentation on applicable TTPs during Focused Dispatch was primarily limited to the Special Text edition of TTPs for a Task Force (U.S. Army Armor School, 1995b), and ancillary TTP documentation for company (U.S. Army Armor School, 1995c) and platoon units. While these provided an invaluable source for general descriptions of the TTPs, they were not based on a comprehensive job analysis with hierarchical delineation of task groups and sequences. These TTP descriptions varied in the amount of detail provided, but were not intended to enumerate the detailed set of responses required at each step in a task sequence. Special TTP Reference Guides (U.S. Army Armor Center, 1995a) were also developed for Focused Dispatch's alpha and beta cases. While these guides provided useful summaries of each case, they did not adequately delineate the specific tasks and conditions associated with the subject TTPs.

Developing and sustaining an adequate task analysis for an AWE may become the responsibility of the expanded functional analysis team. The development of a task analysis during an AWE is a formidable, almost impossible, job. Rather, the expanded team members involved in functional and task analysis should be routinely developing and refining these products before and after AWEs, and then importing and evaluating these products during each AWE. The sustainment of task analyses and related products, such as TSPs and TTPs, is a recurrent issue across AWEs as the subject information technologies are revised and replaced. The AWE's deliberate spiral development of information technologies requires corresponding upgrades for all supporting products, including task analyses.

While it may be prohibitive to develop a formal task analysis of an entire AWE system, the analysis should address the primary tasks associated with the key issues of the evaluation, such as the alpha and beta case TTPs identified for Focused Dispatch. Recommendations for conducting more informal, but directed, analyses of key tasks will be briefly reviewed.

Even informal task analyses require: identification of the conditions required to stimulate task performance, prespecified and timed sequences of task steps, precise delineation of the control actions taken in performance of each task step, determination of the feedback associated with each task step, and an indication of task completion (Meister, 1965). Additionally, users of these task analyses might benefit from analysts' efforts

to identify key decision points for leaders (e.g., when to utilize reserve units, or issue a change in orders), and the conditions triggering these decision points.

Task documentation at this level is needed to inform evaluators' and trainers' understanding of system concept and utilization. Such documentation is also essential for examining the tradeoffs between system variants, such as conventional versus digital TTPs. The potential impact of future information technologies, for example, may be severely compromised by human-computer interfaces that complicate the control actions required for task performance. Comparisons of task analyses for conventional and digital systems would assist in identifying the complementary aspects of each system as well as the modifications in task performance that might be required in the event of degraded digital systems.

Methods for analyzing the cognitive processes underlying military, particularly command and control, tasks are also recommended (Brannick, Prince, Prince & Salas, 1995). Traditional task analyses often focus on discrete, sequential and manual operations, but fail to define the process steps required for more cognitive aspects of a task such as "decide" or "analyze" (Ensing & Knapp, 1995). For their model of command and control tasks at the brigade level, Ensing and Knapp developed a knowledge elicitation method for specifying the mental steps in the thought processes underlying cognitive aspects of task performance. While this work seems especially relevant to the subsequent recommendation on the use of performance models, these authors also suggest that their diagrams of work flow for generic command and control tasks are applicable to TTP evaluation.

In sum, task analysis provides a detailed framework that might support the formative evaluation of key AWE system components such as equipment, operators, and procedures. The ongoing refinement of key task analyses by an expanded team could provide a very useful product to an AWE's evaluation team. More innovative task analysis methods, as reviewed, might help AWE evaluators understand the impact of information technologies on the cognitive processes required for performing key TTPs.

<u>Process Measures</u>

The lack of formative information provided by traditional methods of evaluating training and force capability, accentuates

the need for evaluators to adapt measures more suitable to the AWE objective of improving the force. Training and Evaluation Outlines (T&EO) frequently fail to specify many relevant dimensions of task performance, provide evaluators only cues to "standard" specification, and list tasks in strict chronological sequences that are unrealistic (Ensing & Knapp, 1995; Havron, McFarling & Wanschura, 1979). The restricted format of T&EO items prevents trainers or evaluators from "recapturing" meaningful performance parameters and is "inimical to incisive training diagnostics" (Havron et al., 1979).

The role and importance of process measures in support of formative evaluation objectives were emphasized in a prior section, Formative Evaluation and the AWEs. Advancements in system complexity and soldier-machine interfaces place increased importance on the acquisition and retention of procedural skills (Morrison, 1982). Future AWE evaluators might review more traditional methods for the development of process and procedurally oriented measures (Meister, 1965; Plott et al., 1992). Many of these traditional process measurement methods are reflected in this section's prior AWE method recommendations.

For example, the multidisciplinary makeup of the evaluation team was recommended to provide an accurate and integrated understanding of key equipment, personnel, and procedural elements. Similarly, the focus on process measures presumes that evaluators develop a detailed knowledge of key system functions and tasks. Some of the primary types of process concerns and measurement issues that might be addressed include: the understandability of the process in terms of its completeness and detail including task stimulus, required control actuations and communications, and all necessary feedback; and, the difficulty of the process requirements with respect to coordination, speed, precise discrimination, and input (Meister, 1965).

Explicit Definitions. An emphasis on process measures for key issues and tasks, such as TTPs or the common picture, may guide AWE evaluators in answering the nagging question "What is it we are supposed to measure?". The corollary question "How are we to measure it?" must also be addressed by the evaluation team. Requisite skills of human performance specialists on the evaluation team should include their ability to develop explicit and precise procedures that specify how each selected AWE measure is defined and collected.

Guidelines for defining explicit procedures for military measures are abundant and consistent with prior recommendations for more detailed understanding and documentation of AWE system functions, tasks and procedures. Increased precision in the specification of measurement operations is central to these guidelines (Boorman, 1993). Suggestions for improving observational measures, for example, typify this concern for precision: simplify the observational task; specify as concretely as possible the actual cues the observer should attend and respond to; and, provide training to the observer in cue discrimination and data recording (Meister, 1965).

Future AWE evaluators might review standard taxonomies of military measures, but more pertinent process and intermediate measures are strongly recommended (Brannick et al., 1995; Eddy, 1989; Garlinger & Fallesen, 1988). O'Brien et al. (1992) provide detailed operational definitions of measures for conventional (see also Elliott & Quinkert, 1993) and digital systems for selected battlefield functions, including command and control. These definitions leverage many of the data collection tools available in virtual simulation settings, such as the Mounted Warfare Test Bed used during Desert Hammer VI and Focused Dispatch. An extended bibliography of selected team performance measures that are categorized by outcome measures, process measures, measure selection strategies, and novel measurement techniques is provided by Eddy (1989). Demonstrated methods for evaluating and validating team processes are available and urge, for example, that multiple observations are essential to their accuracy (Brannick et al., 1995).

The relationship of AWE process measures to more traditional measures of outcome is also a concern. Useful recommendations for deriving and linking process oriented measures to traditional outcome measures should help address this issue. For example, a recent handbook for developing command and control MOEs details their relation to process measures and tasks, and also provides a good conceptual overview (Boorman, 1993).

Similarly, Wheaton et al. (1980) describe their use of a front-end mission analysis to identify needed process measures that complemented outcome measures derived from prior analyses. These authors conclude these methods effectively explicated and linked the types of process and outcome performance to be expected, and provided a basis for subsequent generation of performance constructs and measures. Finally, outcome or product

measures might further the AWE's formative purpose by serving as "markers" of effective and ineffective key task performance (Havron et al., 1979). These authors advocate the use of outcome measures (e.g., during After Action Reviews) as probes to investigate underlying processes, and to develop a more valid consensus on criteria for process measurement.

In sum, once program goals and research issues are more precisely defined and evaluators have a solid understanding of key system functions and tasks, the specification of evaluation measures is relatively straightforward. The overall set of method recommendations and guidelines herein should provide a useful foundation for developing AWE measures, particularly, process measures. Early formulation of explicit definitions allows evaluators, and directors, to realistically assess each measure against available data collection resources for each AWE research setting. Overall, these assessments should guide allocation of research issues across the AWEs and related RDA&Tng efforts, support the use of qualitative measures where necessary, and accent the need for automated instrumentation.

Quantitative and Qualitative Measures. Directors and evaluators should be realistic about the type and quality of data obtainable during macro-level portions of AWEs. Recall that the precision of the resultant test data is dictated by the stage of system development and Force XXI information technologies are in the early stage of development. Paradoxically perhaps, a shift to more formative AWE assessment methods enhances prospects for meaningful quantification, for obtaining reliable and valid measures (Dewar et al., 1994; Sackett & Larson, 1990). A focus on more intermediate goals and measures reduces the scope of the evaluation, shortens the chain of complexity, and provides substantially more data points from an exercise than the "one per run" often obtained with outcome MOEs.

Moreover, the need and potential utility of qualitative AWE indicators for improving key aspects of equipment, procedure and training for future forces should not be underestimated. To develop a reliable framework for more precise quantitative and qualitative AWE process measures, prior recommendations have stressed the role of functional and task analyses of TTPs for explicit descriptions of the discrete behaviors required for performance.

For more procedural TTPs, such as "sensor to shooter" fire requests and dispersed movement, quantifiable performance measures should be relatively easy to compile given the overt nature of performance. Less procedural tasks, however, such as "decide" and "analyze," may require that subject matter experts initially decompose such tasks into more meaningful and measurable steps.

Useful methods for deriving and structuring such performance into more discrete procedural steps are documented (Ensing & Knapp, 1995; Brannick et al., 1995). Qualitative and subjective measures may be needed to assess less procedural aspects of TTP performance, such as cognitive and team-based performance. Relevant methods for collecting expert ratings on team process dimensions such as decision making, communication, and situational awareness in a military context are recommended (Brannick et al., 1995; Burnside, 1982). For an extensive review of other methods for eliciting the types of knowledge and processes used in complex decision making, see Cooke (1994).

Useful AWE qualitative measures include: participants' explanations of why certain performance did or did not occur; analyses of why errors or failures occurred; descriptions of interface and function inadequacies; and, the attitudes of participants and others to test situations, functional allocations, TTPs and training. The most meaningful data obtained on comprehensive system-level evaluations is often qualitative and subjective (Meister, 1987). In the AWE effort to formatively improve force capability, the explanatory power of qualitative and subjective measures should be instrumental to modernization objectives.

Instrumentation. The directed use of automated performance measurement devices is strongly recommended for AWE efforts. In the context of the AWEs, automated performance recording might greatly assist capturing the density of data on system outputs available, and substantially reduce the workload imposed on human data collectors (U.S. Army Armor Center, 1995a). In particular, the continuous and cross-unit nature of many AWE measurement issues, such as maintaining a common picture of the battlefield, accentuates their candidacy for automated measurement.

More common forms of automated data recording, such as audio and video records, can faithfully and continuously capture records of performance that can be later examined. Such records

contain a volume and detail of information that far exceeds the capabilities of human data collectors. While the extraction and reduction of such information can be costly and time consuming, considerable efficiency is gained when precisely scoped issues and measures are combined with the time-stamped flagging of key events in structured scenarios (Leibrecht et al., 1994). The resources required for recording and reducing such data, however, are substantial.

Past AWE evaluation efforts were limited by noninstrumented information systems (U.S. General Accounting Office, 1995). The digital information systems, integral to the AWEs, were like black boxes that provided no record of operator actuation such as the responses taken in acquiring, processing and disseminating information. The incompatibility of these different AWE communication systems, frequently forced operators to manually reenter and then relay messages from one system to the next, swivel-chair integration. This incompatibility precluded attempts to automatically record the flow of information between participants using different information systems.

Ironically, the resources required for instrumenting advanced information systems are almost inherent to their computer-based nature. Instrumentation costs for such systems, however, should be amortized across AWEs and related Army efforts. Evaluators might champion the potential of instrumentation for answering important AWE issues, including many aspects of TTP performance and the common picture process.

Performance Models

Efforts to develop a detailed understanding of the TTPs for employing information technologies might benefit by developing models of TTP performance. Notably, many important TTPs entail higher-order cognitive processing in multitask environments, including: troop leading procedures for digitally-equipped units; requirements to gather and analyze information from a variety of information devices for intelligence operations; and, the requirement to restructure conventionally linear and sequential tasks, such as planning, into parallel tasks performed concurrently by multiple participants. Direct measures of observable participant performance may not adequately address the covert cognitive aspects of such performance. Models of human performance, however, might supplement direct measures and guide

development of the formative data bases needed to improve digital TTP performance.

Performance models of human-computer interaction can support performance predictions for subject systems by combining system specifications and structural constraints for a given task with values associated with the human performance of that task. System developers should provide the structural specifications for their system. Human performance values for each task may be based on actual human inputs, mathematically-based logical inputs, or extracted from premeasured or precalculated data bases (Card, Moran & Newell, 1983).

Performance models are useful for: predicting the effect of proposed engineering changes prior to costly revisions; predicting the operational effectiveness of the system; indicating where improvements in system performance are required; identifying critical links between equipment, personnel and the sequence of mission events; and, providing human factors inputs to tradeoff decisions (Meister, 1965). With the exception of operational effectiveness, these uses of performance models are all central to the formative development of information technologies and their impact on TTPs. As with the function and task analyses upon which these models are linked, the concurrent availability of a conventional performance model of key TTP performance might identify the complementary aspects of each system as well as system tradeoffs.

More recent examples of military investment in performance models for the management of complex systems are cited by Adams, Tenney and Pew (1995). These authors strongly endorse the use of performance models for complex systems as: powerful tools for hypothesis generation; "live" databases for checking hypotheses through simulation; theoretical basis for understanding the interrelationships of humans and equipment; and, ultimately for simulating human performance in the design of better systems.

To illustrate the potential utility of such models to the AWEs, two empirical examples of performance model effects are briefly reviewed. In an evaluation of current versus proposed interface designs for tollbooth operators, John (1995) compared an analytic prediction from a performance model against an independent field evaluation. Interface designers predicted that the new interface would result in a 20% reduction in operator time, and a savings of $3 million annually. John's performance

model, however, predicted operators would perform 3% slower, and this prediction was precisely confirmed in the subsequent field evaluation. The model also provided an explanation of this finding: interface changes impacted subcomponents of the design that were not germane to an operator's time-line for critical tasks.

Similarly, in a constructive simulation of digital versus conventional brigades, the U.S. Army Armor Center (1993) used a commercial modeling language to insert tactical communication values (e.g., message frequency, time to transmit, and process messages) into a battle command network. Although based on an optimal model of communication linkages, the observed trends indicated that information technologies save time in moving information within and between units, and that units with digital systems can react more swiftly to battlefield events. This evaluation's design compared identical moments in the battles, based on the reaction times associated with digital versus conventional information systems.

One notable result of this modeling approach addressed the unit's detection and reaction to an enemy counterattack and found that "the digitized force was executing three decision cycles to the enemy's one" (U.S. Army Armor Center, 1993). While such optimal-model results may exceed those obtained in more realistic live or virtual simulation, they provide an intriguing indication of the force capability improvements possible when more capable and robust information technologies are evaluated.

The resource requirements for performance models are also substantial and should be amortized across AWEs and related Army efforts. Justification of these resources might stress the potential of performance models to make valuable contributions to the development and refinement of complex human-machine systems, and their role in integrating the Army's related research and development efforts. The military's prior investments in support of related performance models should also be capitalized by the AWEs. For example, classic examples of command and control performance models (Baker, 1970; Siegel & Wolf, 1969) provide a useful basis for guiding the development and adaptation of such models to the AWEs. More recently, the command and control performance model described by Ensing and Knapp (1995) was developed for a prototype Command and Control Vehicle (C^2V), a system integral to many AWEs.

The task of actually developing the performance models applicable to the AWEs is beyond the role of the core evaluation team. Development and iterative refinement of such models requires an expanded team of modeling experts who provide periodic service to the core evaluation team. Information laboratories, such as the Communications and Electronic Command (CECOM) and its Digital Integration Lab, might effectively integrate the Army's performance modeling efforts across BLWEs, ACTDs, ATDs and related research and development.

Performance modeling efforts from each of these areas, as well as more traditional "bench" tests of breadboard and brassboard systems, should directly support the AWEs. Such AWE-related efforts might: verify the espoused functionality of systems prior to an AWE; corroborate the potential deltas of systems in controlled laboratory settings; identify selected research issues for an AWE; and refine their models based on AWE findings. As discussed earlier, the more robust results available from such efforts should contribute to Joint Venture's supporting body of evidence.

Structured Scenarios

The relatively unstructured nature of extended and large-scale combat operations is a major impediment to military evaluation. More general difficulties in conducting military evaluations, particularly in field or live simulation settings, were previously examined. Researchers, and even directors, are unable to control the multitude of extraneous variables associated with complex, field-based operations. The scope and complexity of the AWEs progressively compound evaluation difficulties by increasing the size of the test unit, and by iteratively modifying equipment, personnel and procedural elements over the course of the AWEs. Even the AWEs reliance on multiple settings--live, virtual and constructive--seriously complicates controlling and measuring the "same" variables in different evaluation settings.

Fortunately, the Army is making substantial progress in the use of virtual and constructive simulation to structure training and evaluation activities. This progress should be leveraged in the development of AWE operational scenarios and supporting TSPs. The current capstone method for structuring combat scenarios is the Virtual Training Program (VTP) developed using virtual and constructive simulations at Fort Knox (Campbell, Campbell,

Sanders, Flynn, & Myers, 1995). The VTP methodology supports the development of complete training support packages for multiechelon and collective combat operations. In particular, this methodology is grounded by a set of structured and tactically realistic scenarios.

The VTP method and documentation provide evaluators instrumental guidance on the development of critical task lists, such as tasks associated with TTPs, and the sequencing of tasks within each scenario. Important considerations in determining the sequence of tasks within each scenario are also addressed, such as crawl-walk-run, natural or chronological order, and easy-to-difficult. Central to the VTP's structure is the detailed specification of scenario factors such as the location, organization, status and disposition of friendly and enemy forces, and the delineation of scripted events during the exercise.

Overall, the methods developed for the VTP provide trainers and evaluators a systematic and comprehensive set of guidelines to better ensure key tasks are structured and tailored to the constraints and capabilities of designated virtual and constructive settings. Standard VTP scenarios were used during early phases of Focused Dispatch to train participants on fundamental combat operations with conventional systems in virtual simulation (U.S. Army Armor Center, 1995a). For future AWEs, evaluators might adapt a digital version of the VTP scenarios, recently developed (Winsch, Garth, Ainslie, & Castleberry, 1996).

From an evaluation perspective, even when decomposed into a subset of mission segments called "tables," the VTP scenarios may favor tactical realism for training over standardized control for evaluation. For example, the VTP methodology recommends that scenarios should be designed to maintain a continuous flow of battlefield events that unfold across contiguous terrain. Such recommendations are consistent with the VTP objective, realistic training. Evaluation objectives, however, may require more standardized battlefield conditions and the replication of those conditions for repeated practice and measurement. Nevertheless, the pragmatic scope of the AWEs underscores their need for realistic, full-mission scenarios.

More controlled operational conditions are not precluded by the AWEs, however, and would support the fundamental need to

develop reliable data and information bases. Future AWE evaluators might supplement full-mission scenarios and findings by also conducting directed evaluations with more restrictive conditions, as in true subexperiments. General strategies for increasing control in military testing are recommended (Boldovici & Bessemer, 1994; Meister, 1965), and specific examples of such controls in each simulation setting are briefly noted.

For field-based settings, excellent examples of methods for evaluating platoon battle runs are provided by Wheaton et al. (1980). For constructive settings, evaluators might review the modeling effort that focused on identical moments in the battles of digital and conventional brigades (U.S. Army Armor Center, 1993). In virtual settings, a variety of tools and techniques are available to efficiently standardize battlefield conditions (Atwood, Winsch, Quinkert, & Heiden, 1994). Empirical examples utilizing these tools include a series of structured data collection exercises (DCEs) for armor battalions (Lickteig & Collins, 1995), and standardized vignettes to assess information processing performance (Lickteig & Emery, 1994).

In sum, the Army has and continues to make impressive strides in the development of structured training programs and packages (U.S. Department of the Army, 1996c) to better prepare soldiers and units for the operational requirements targeted by an AWE. Future AWE trainers and evaluators should note that the VTP efforts have focused primarily on the execution phase of operations, and additional work may be needed to apply these methods to other operational phases, such as planning and preparation. The VTP methods exemplify the Army's effort to exploit virtual and constructive simulation for structured training. The AWEs should leverage and extend such methods for structuring both full-scale scenarios and more restricted exercises for subexperiments.

Functional Tests

Shortcomings in developmental equipment reinforce the vexing observation: "In the Army 'waiting' is an intransitive verb, there is no object." Such shortcomings have plagued the AWEs (U.S. Army Armor Center, 1994, 1995a) and confounded the results obtained. Attempts to redress such shortcomings, such as cutoff dates for "good ideas" and "freezes" on software upgrades, have had only limited success. The scope and tempo of the AWEs suggest that no panacea to such recurrent shortcomings is likely.

Recall, the pragmatic nature of the AWEs was a key consideration in the proposal for formative AWE evaluations. Nevertheless, AWE evaluators have not made use of the power of functional tests, sometimes called acceptability or developmental tests, to: proactively identify such shortcomings, spur and direct efforts to overcome them, and document system status during training and evaluation.

Guidelines for the development of functional tests are available (Meister, 1965; Plott et al., 1992) and should provide a basis for adaptation to the AWEs. Also, a detailed functional test for an evaluation in virtual simulation was developed by Heiden, Sever, Smith and Throne (1996). This test addressed functional requirements for a digital command and control system, similar to those tested in the AWEs, and might be readily adapted to AWE virtual simulation efforts. As documented by these guidelines, the functional test plan should specify test procedures and criteria that are commensurate with the key functionality required for evaluation, such as TTPs. In general, a functional test should address required functionality of the test system and the supporting test setting.

In a virtual setting, for example, supporting functionality tests may begin with the simulated weapon systems and their relevant component subsystems, such as advanced information technologies. Such tests might then extend to the communication architectures linking all distributed simulators and their information-based systems. In turn, the anticipated functionality of all semiautomated friendly and enemy units required for AWE exercises should be assessed. Additionally, these functional tests should include the virtual test bed utilities for initiating, controlling, recording and analyzing the AWE's training and evaluation exercises. Sampling strategies may be needed to provide a comprehensive test, but should include a thorough check on functionality required for key task performance, such as the TTPs. Initial test phases might sequentially address functions at each unique workstation; but later tests should verify key functionality in a fully-loaded operational setting, all systems operating simultaneously.

In sum, predicated functionality of subject technologies and the overall system is critical to the success of the AWEs. Pervasive and systemic dysfunctions that prevent or disrupt task performance frequently lead to meaningless and/or negative results on outcome MOEs and MOPs for summative evaluations. Even

for formative evaluations, shortcomings in the functionality of the test system may seriously restrict and contaminate training and evaluation efforts. AWE evaluators should carefully document what works and what doesn't to inform the AWEs' conclusions, sponsors, and critics.

Trained Participants

Trained AWE participants are paramount to AWE efforts, exceeding even the need for full system functionality: soldier skills should ultimately include the ability to overcome equipment deficiencies. The AWE's introduction of developmental information technologies into combat units impose a challenging set of new training requirements. Past AWE efforts suggest that information technologies create a progressive hierarchy of training criteria: combat fundamentals, digital fundamentals, and their integrated application in simulated combat, as indicated in Figure 5 (U.S. Army Armor Center, 1994).

Many of the TTPs for employing digital systems may equate to the highest level of that hierarchy. For example, the evolving tactics and techniques that fully leverage the potential of information systems, such as parallel planning, often presume mastery of combat and digital fundamentals. Mastery training is very important when performance deltas are hypothesized. It also provides the skills and confidence required for conducting successful operations in demanding and prominent military exercises, such as AWEs conducted at the National Training Center (J. R. Witsken, personal communication, 22 March 1996).

Although digital TTPs are critical to every AWE (Rigby, 1995), training programs and documentation for digital TTPs are minimal (Quinkert & Black, 1994). A fundamental challenge to the AWEs, therefore, may be the development of detailed Training Support Packages (TSPs) for a digitally-equipped force. The scope of that training development challenge complements Joint Venture's force development effort, spearheaded by the AWEs. The complementary nature of these developmental efforts should be reflected in every AWE, as they were in Focused Dispatch, with its primary goals of TTP refinement and the development of a TSP for a digitized Task Force (U.S. Army Armor Center, 1995a). AWE evaluators should carefully attend to the lessons learned about AWE training (Elliott, Sanders & Quinkert, 1996; Kollhoff, 1995).

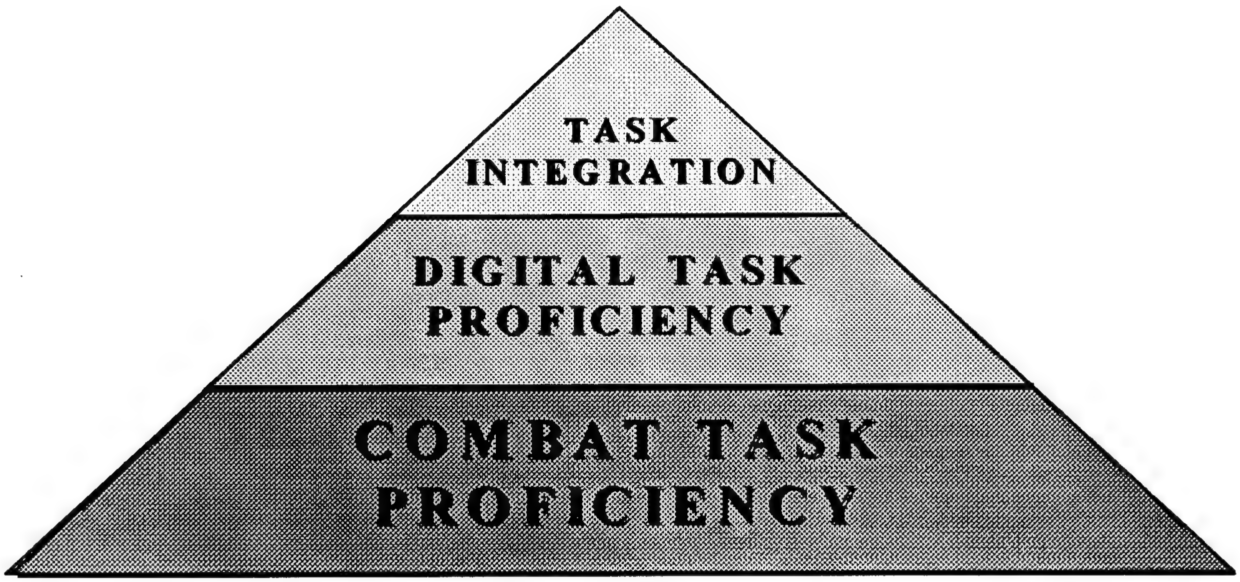

Figure 5. Hierarchy of skills and training for achieving digital warfighting capability (Adapted from U.S. Army Armor Center, 1994).

The importance of adequate TTP training was highlighted by the work of Root, Hayes, Word, Shriver and Griffin (1979). Their evaluation of field-based techniques for training tactics, demonstrated that simply "dumping" new tactical techniques on participants was insufficient. They found that members of their evaluation team frequently had to step-in and teach necessary procedures. One of their major recommendations was that officers and noncommissioned officers "...should not be expected to read and digest bulky documentation when much of it represents significant changes to existing training procedures" (p. 34). These authors concluded that the success of new tactics requires assimilating unfamiliar concepts with new rules and techniques, and that their interrelationship are hard to visualize (see also U.S. Army Armor Center, 1994).

The TTP proficiency of the participants should be the primary concern of AWE evaluators assessing TTP performance. For that reason, a trainer was included in the core evaluation team to establish concurrence on training and evaluative objectives for key TTP aspects of performance. Early concurrence is crucial to the success of the AWE, and exemplifies the interdependent roles and needs of AWE trainers and evaluators that permeate all phases of a formative evaluation. Their common need for a detailed understanding of the system based on function and task analyses, for example, highlights the requirement that evaluators and trainers continuously share and coordinate these products.

Training and evaluation scenarios should structure operations that systematically require performance of the key tasks and TTPs, their complementary objective.

In the future, AWE evaluators might adapt useful guidelines for evaluating training to ensure participants can proficiently perform key AWE tasks, particularly key TTPs (Kristiansen and Witmer, 1981). Trainers and evaluators should push for a structured training support package that progresses participants through the fundamentals of combat and digital operations, and then their integrated application in realistic combat situations. A primary purpose of most AWEs is to gather and organize information that formatively improves training, an imperative to improving the force.

Trained Data Collectors

Data collection is the essential element in all evaluation. Data collectors are the "front-line" of the evaluation team and their proficiency is a primary determinant of data quality. Many of the prior method recommendations bear directly on developing an AWE evaluation "environment" supportive of data collection requirements. For example, it was recommended that the primary requisite for leadership of the multidisciplinary evaluation team should be expertise in developing precise procedures that specify how each selected AWE measure is defined and collected. Similarly, data collectors should comprise an expanded evaluation team with early and routine representation. Before considering additional recommendations, a brief review of the data collector training problems that affected Focused Dispatch should highlight key concerns.

The primary data collectors for Focused Dispatch were called the EXFAC to approximate their myriad roles as Experimental Facilitators. Their roles were integral to that AWE effort and included directing, observing and controlling the conduct of training and evaluation exercises in live, virtual and constructive settings. In addition, the EXFAC organized and conducted the After Action Reviews (AARs) for Focused Dispatch, including preliminary and culminating AARs associated with each mission exercise (U.S. Army Armor Center, 1995a). As observers, raters and data collectors, the EXFAC administered most, and completed many, of the manual data collection instruments used in Focused Dispatch. The inherent conflict in their multiple roles and responsibilities was compounded by the fact that the EXFAC

were belated members of this AWE's evaluation team, and not routinely included in evaluators' early meetings on the purpose and scope of Focused Dispatch. The EXFAC also had limited involvement in many of the subsequent evaluation meetings directed at identifying measures, instruments and data collection procedures.

Initially, these EXFAC were a conventionally, versus digitally, skilled group of military subject matter experts. They routinely served as dedicated trainers for the Virtual Training Program (VTP) exercises conducted at Fort Knox. Such expertise was instrumental to their many roles for Focused Dispatch, including their conduct of several conventional VTP exercises used to ensure the AWE participants were proficient in combat fundamentals (Elliott, Sanders & Quinkert, 1996).

The EXFAC's conventional expertise was later diminished, however, as their AWE assignments were not stabilized during the course of Focused Dispatch. Many of the more experienced EXFAC members were reassigned and replaced with newer personnel during this AWE. With respect to digital expertise, however, the EXFAC had very limited hands-on experience with the AWE's digital information systems. Moreover, the EXFAC had no formal training on these digital systems or on the TTPs for their employment. Absent documentation, such as functional and task analyses, limited the EXFAC's options for developing a more detailed understanding of these systems.

The AWE's macro-level complexity and developmental systems exacerbated the EXFAC's workload and blunted the precision needed to specify measurement tools and procedures. Operational vagueness in measures, such as the number of "substantive" and "useful" messages, frustrated the EXFAC's data entry and collection efforts. The lack of instrumentation for digital systems, particularly for capturing data on operator actuations and information flow between different systems, resulted in excessive workloads on these data collectors. The failure to systematically embed key tasks, such as TTPs, into structured scenarios and data collection instruments, severely complicated attempts to provide these data collectors with sequentially-ordered measures and predictable cues.

Data quality concerns urge that such data collection problems, as identified for Focused Dispatch, are corrected for future evaluations. Initially, the core evaluation team should

insist on early and frequent coordination meetings with key data collectors that address evaluation issues and measures in terms of data collection resources and procedures. As the expanded team of data collectors is formed, evaluators should strive to stabilize membership to maintain data collectors' expertise throughout the course of the evaluation.

Evaluators should marshal available documentation and equipment to provide all data collectors a clear understanding of AWE system functions, and hands-on operator experience of the tasks slated for evaluation. After data collection procedures and activities are defined, the core team should conduct general familiarization and detailed hands-on training sessions with all data collectors to rehearse and refine data collection procedures and instruments. The schedule for these training sessions should allow sufficient time for revising procedures and instruments based on data collectors' feedback.

The AWE's core evaluation team should also develop efficient and effective data collection procedures by applying guidelines and methods for structuring and simplifying data collection requirements. Useful methods for increasing the precision of the data collection task include: simplify the observational task; specify as concretely as possible the actual cues the observer should attend and respond to; and, provide training to the observer in cue discrimination and data recording (Meister, 1965). AWE evaluators might also adapt available job aids for structuring observational requirements (Witmer, 1981).

In the live portions of an AWE, particularly the high-visibility rotations at the National Training Center, the training, experience and workload of the data collectors assumes pointed importance. During the 1994 AWE, Desert Hammer VI, the dedicated Observer/Controller team at this location provided invaluable assistance in collecting and providing data. These O/Cs primarily used 5-point checklists organized by Battlefield Operating System and unit level. Despite the checklists' relatively simple and compressed format, the additional requirement to monitor and assess digital performance severely burdened these skilled data collectors (J. R. Witsken, personal communication, 22 March 1996). Nevertheless, the broad-based experience of these O/Cs remains an excellent resource for future AWEs.

In sum, future AWE evaluators should insist that at least some key members of the expanded data collection team are consulted and involved in all data-related phases of an AWE evaluation. The formal and informal training of data collectors should ensure they all have a clear understanding of system capabilities and functions, operator and equipment performance requirements, and the purpose and requirements of data collection.

Utilization of Findings

For AWE efforts directed at improving force capability, an acceptable exit criterion might be the implementation of lessons learned. This implementation criterion follows from the premise that the primary objective of Joint Venture and the AWEs is improving force capability. The methods of formative evaluation applied to the AWEs should provide the exploratory and explanatory power required for learning many of the lessons essential to improving force capability. The logical, but not always achieved, objective of learning lessons is to use them.

A primary component of a formative AWE rolling baseline that directly contributes to the process of improvement might be "living products" that iteratively implement lessons learned across the force and training development spectrum of Joint Venture. The term living products, versus living documents, is used to underscore the utility and the DTLOMS-wide range of Force XXI lesson implementation.

Living product examples include TTP Special Texts and TSPs, operational requirements, system specifications, software applications and communication protocols for information-based technologies, performance models, process-related findings, and an Evaluation Support Package. Product revisions should be based on a Model-Evaluate-Model-Evaluate (MEME) approach applied across the AWEs and Army-wide Research, Development, Acquisition and Training (RDA&Tng) efforts (see Figure 6). Notably, this MEME model does not address the issue of validation, given this report's emphasis on formative evaluation and force capability improvement. The MEME model underscores collaborative product refinement and utilization to ensure lessons learned are implemented across Army-wide efforts to achieve Force XXI capabilities. In effect, such living products might provide a common-picture/product synergy to Joint Venture efforts.

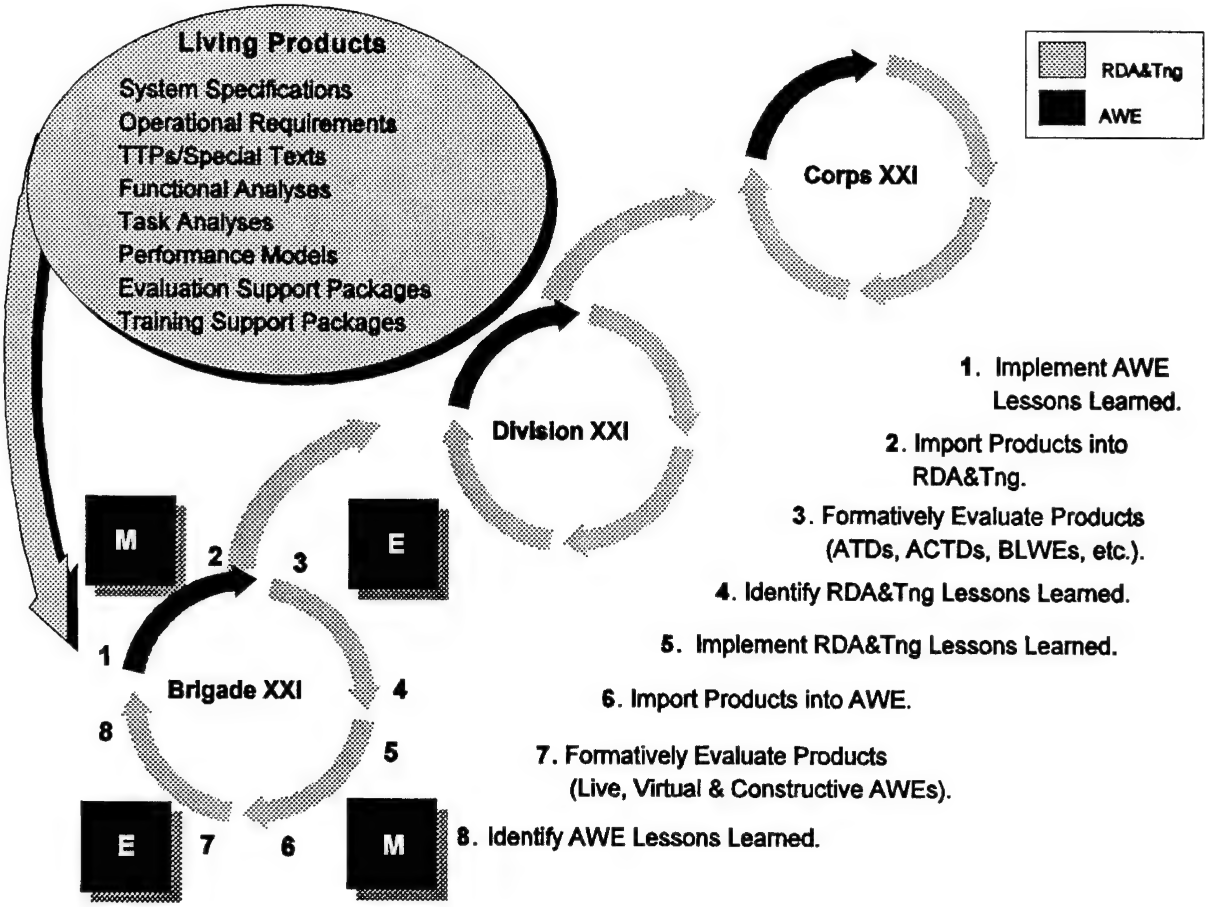

Figure 6. Living product development cycle based on a Model-Evaluate-Model-Evaluate (MEME) approach to lesson implementation across AWEs and RDA&Tng.

Lesson implementation is often more difficult than lesson learning, and more useful than lesson documentation. The AWEs have already compiled many valuable lessons learned that should improve future force capability. Desert Hammer VI, for example documented lessons across DTLOMS (U.S. Army Armor Center, 1994) with particular emphasis on doctrine (Witsken, 1995), training (Kollhoff, 1995) and materiel (Vowels, 1995). Lessons learned from Focused Dispatch are documented in the MBBL's forthcoming After Action Review of that AWE, with emphasis on the digital TTPs underlying doctrine and a TSP for training digitally-based operations (see Elliott, Sanders & Quinkert, 1996). The mere documentation of lessons learned, however, often results in relearning and redocumenting the same lessons.

Notably, the primary Focused Dispatch deliverables stressed the implementation of lessons learned as an exit criterion for

this AWE (see Figure 3). Training lessons from this AWE were implemented into a set of TTP documents, as part of a TSP, for delivery to the Experimental Force that will serve as the participant unit for Task Force XXI. The TTP lessons learned from this AWE form a basis for revising the Special Text TTPs for digitally-equipped brigade, battalion, and company units (U.S. Army Armor School, 1995a-c). Despite shortcomings in this AWE, these two AWE deliverables exemplify the goal of implementing lessons learned into living products.

In sum, implementing lessons for Force XXI capabilities requires a collaborative mechanism that iteratively imports and exports lessons learned across the Army's force development and training efforts. Embedded in this report's method issues and recommendations is a mechanism for lesson implementation, the expanded evaluation teams. Beginning with the formation of a multidisciplinary evaluation team and its expanded teams, these method recommendations stress that the AWEs are not stand-alone evaluations (see Figure 1). The expanded teams, drawn from related Army agencies and programs, should empower the AWEs by importing their respective products into the AWEs. They should supplement the AWEs by revising their products based on lessons learned during the AWE and exporting those products back to their respective agencies and programs. Expanded teams are a network mechanism that links the Army's Force XXI efforts. Living products are an explicit medium for lesson implementation leading to force improvement.

Summary and Conclusion

The U.S. Army's venture toward future capabilities is spearheaded by the Joint Venture Campaign Plan to redesign and develop the operational force. A cornerstone of this plan is a series of AWEs to iteratively discover how the Army should equip, train and fight the future force, Army XXI. This report focuses on the Army's Force XXI Advanced Warfighting Experiments (AWEs) supporting this objective, particularly research methods appropriate to the AWEs and related Army modernization efforts.

The Introduction section reviews how the broad scope and purpose of the AWEs pose a serious challenge to military researchers and more traditional approaches to military testing. This review stresses that the primary objective of Joint Venture and the AWEs is to improve force capability, and a preliminary concern with formative issues and methods may avoid summative

conclusions of failure. In support of that objective, formative research methods address process and intermediate measures that enable more final objectives, such as "proving" improved force capability.

The Method section identifies and then addresses twelve fundamental and formative evaluation issues for the AWEs. The AWE formative evaluation methods recommended herein stress the need for research methods that provide solutions, rather than identify failures. The AWE fundamental evaluation methods recommended underscore the principles of "good" research (Thompson and Rath, 1974) for the AWEs.

The Utilization of Findings section suggests that an acceptable exit criterion for the AWEs, as formative evaluations, might be the implementation of lessons learned. The primary findings of an AWE that directly contribute to the process of improvement might be "living products" that iteratively implement lessons learned across the force and training development spectrum of Joint Venture. Product revisions might be based on a Model-Evaluate-Model-Evaluate (MEME) approach applied across the AWEs and Army-wide RDA&Tng efforts. These living products might provide a common-picture/product synergy to Joint Venture efforts.

Method recommendations throughout the report embed a mechanism for implementing AWE findings into living products. Expanded AWE evaluation teams should import their respective products into the AWEs, revise their products based on lessons learned during the AWE, and then export those products back to their respective agencies and programs. These expanded teams could link the Army's Force XXI efforts, and their living products could provide a medium for Army-wide force improvement.

In closing, the Joint Venture Campaign Plan stresses that the Force XXI pathway to the Army's future is a collaborative and iterative developmental process to achieve an objective, Army XXI. This report attempts to support that objective by focusing research methods on the process of achievement.

References

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. Human Factors, 37(1), 85-104.

Army. (1994, October). The end of the beginning. Army, 59-66.

Army. (1995, June). A talk with the chief. Army, 14-24.

Atwood, N. K., Winsch, b. J., Quinkert, K. a., & Heiden, C. K. (1994). Catalog of training tools for use in distributed interactive simulation (DIS) environments (ARI Research Product 94-11). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A282 841).

Baker, J. D. (1970). Quantitative modeling of human performance in information systems. Ergonomics, 13, 645-664.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). Handbook on formative and summative evaluation of student learning. New York: McGraw Hill.

Boldovici, J. A. (1995). Statistical power and generality in training-device evaluations. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Boldovici, J. A., & Bessemer, D. W. (1994). Training research with distributed interactive simulations: Lessons learned from simulation networking (ARI Technical Report 1006). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A285 584)

Boorman, L. G. (1993). Command and control measures of effectiveness handbook (C2MOE Handbook). (TRAC-TD-0393). Fort Leavenworth, KS: TRADOC Analysis Command-Study and Analysis Center.

Bouchard, T. J., Jr. (1976). Field research methods: Interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures. In M. D. Dunnette (Ed.), Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally.

Brannick, M. T., Prince, A., Prince, C., & Salas, E. (1995). The measurement of team process. Human Factors, 37(3), 641-651.

Burnside, B. L. (1982). Subjective appraisal as a feedback tool (ARI Technical Report 604). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A138 873)

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi experimental designs for research. Chicago: Rand McNally.

Campbell, H., Campbell, R. C., Sanders, J. J., Flynn, M. R., & Myers, W. E. (1995). Methodology for the development of structured simulation-based training (ARI Research Product 95-08). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A296 171)

Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.

Colwell, V. J. (1993, October-December). Battle command battle lab--Fort Gordon. Military Intelligence, 9-10.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. International Journal of Human-Computer Studies, 41, 801-849.

Cosby, N. L. (1995). SIMNET: An insider's perspective (IDA Document D1661). Alexandria, VA: Institute for Defense Analysis.

Cronbach, L. (1964). Evaluation for course reform. In R. Heath (Ed.), New Curricula. new York: Harper & Row.

Dewar, J., Bankes, S., Hodges, J., Lucas, T., Saunders-Newton, D., & Vye, P. (1994). Credible uses of the distributed interactive simulation (DIS) system (RAND DRR-878-1-A). Santa Monica, CA: RAND.

Dubin, H. (1995, February). Applying the Force XXI baseline assessment strategy (Brief to General Officer Working Group). [Slides]. (Available from Operational Test and Evaluation Command, 4501 Ford Avenue, Alexandria VA 23651).

Eddy, D. R. (1989). Selected team performance measures in a C3 environment: An annotated bibliography (USAFSAM-TR-87-25). Fort Detrick, MD: U.S. Army Medical Research and Development Command. (AD A210 695)

Elliott, G. S., & Quinkert, K. A. (1993). Performance measures for simulation (ARI Research Product 93-08). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A267 837)

Elliott, G. S., Sanders, W. R., & Quinkert, K. A. (1996). Training in a digitized battalion task force: Lessons learned and implications for future training (ARI Research Report 1695). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Ensing, A. R., & Knapp, B. G. (1995). Modeling Army C2 at the brigade level. Paper presented at 1995 International Symposium of Command and Control Research and Technology, Washington, DC.

Fitz-Gibbon, C. T., & Morris, L. L. (1987). How to design a program evaluation. Beverly Hills, CA: Sage Publications.

Garlinger, D. K., & Fallesen, J. J. (1988). Review of command group training measurement methods (ARI Technical Report 798). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gregory, R. (1928). Discovery: Or the spirit and service of science. New York: Macmillan.

Havron, M. D., McFarling, L. H., & Wanschura, R. G. (1979). Improved ARTEP methods for unit evaluation Volume VI: Conventional ARTEP missions and engagement simulations: An examination of options (ARI Technical Report TR-79-A24). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Heiden, C. K., Sever, B., Smith, P. G., & Throne, T. (in preparation). Functional test example for distributed interactive simulation (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. Manuscript in preparation.

Hewish, M. (1995, November). Advanced warfighting experiments. International Defense Review, 31.

Hollis, W. W. (1995, March). Walt Hollis on Army OR: "The times they are a-changin!". Phalanx: The Bulletin of Military Operations Research, 28(1), 1-8.

John, B. (1995). Applying psychology to the design of computer systems. CSERIAC Gateway, 6(3), 7-8.

Johnson, E. M., & Baker, J. D. (1974). Field testing: The delicate compromise. Human Factors, 16, 203-214.

Kollhoff, R. K. (1995). The lessons of Operation Desert Hammer VI: Digitization will impact many areas of training. Armor, September-October, 41-43.

Kristiansen, D. M., & Witmer, B. G. (1981). Guidelines for conducting a training program evaluation (TPE) (ARI Research Product 81-18). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A120 775)

Langford, G. B. (1995). Teaming for integrated product and process management. Army Research, Development and Acquisition Bulletin, 7-10.

Leibrecht, B. C., Meade, G. A., Schmidt, J. H., Doherty, W. J., & Lickteig, C. W. (1994). Evaluation of the combat vehicle command and control system: Operational effectiveness of an armor battalion (ARI Technical Report 998). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A282 767)

Lickteig, C. W., & Collins III, J. W. (1995). Combat vehicle command and control system evaluation: Vertical integration of an armor battalion (ARI Technical Report 1021). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A292 718)

Lickteig, C. W., & Emery, C. D. (1994). Information management performance of future platoon leaders: An initial investigation (ARI Technical Report 1000). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A282 908)

Mann, J. (1972). The outcome of evaluative research. In C. H. Weiss (Ed.), Evaluating action programs: Readings in social action and education (pp. 265-282). MA: Allyn & Bacon.

Martin, W. C. (1995, May -June). Training in the 21st century--the Force XXI training program. Armor, 40-41.

McCall, M. W., & Bobko, P. (1990). Research methods in the service of discovery. In M. D. Dunnette & L. M. Hough (eds.), Handbook of industrial and organizational psychology (2nd ed.) (Vol. 1, Section II, Chap. 7, pp. 381-418). Palo Alto, CA: Consulting Psychologists Press.

Meister, D. (1965). Human factors evaluation in system development. New York: Wiley.

Merritt, J. N. (1995). A talk with the chief. Army, 5(6), 14-24.

Mitroff, I. I. (1985). Why our old pictures of the world do not work anymore. In E. E. Lawler III, A. M. Mohrman, S. A. Mohrman, G. E. Ledford, Jr., and T. G. Cummings (Eds.), Doing research that is useful in theory and practice (pp. 18-43). San Francisco: Jossey-Bass.

Morrison, J. E. (1982). Retention of armor procedures: A structural analysis (ARI Technical Report 591). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A140 003).

Nadler, D. A., Shaw, R. B., Walton, A. E., & Associates. (1995). Discontinuous change. San Francisco: Jossey-Bass.

O'Brien, L. H., Wigginton, D., Morey, J. C., Leibrecht, B. C., Ainslie, F. M., & Sawyer, A. R. (1992). Data tables from combat vehicle command and control battalion-level preliminary evaluation (ARI Research Note 92-84). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A258 580)

O'Bryon, J. F. (1995, July-August). Test, evaluation community responds to new challenge. National Defense, 64-65.

Plott, c. C., LaVine, N. D., Smart, D. L., & Williams, G. s. (1992). Concept analysis for simulation modification methodology (ARI Research Report 1613). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A250 947)

Quinkert, K. A., & Black, B. A. (1994, November-December). Training for Force XXI technologies. Army Research, Development and Acquisition Bulletin, 44-46.

Rigby, J. W. (1995, November). The U.S. International digitization strategy. International Defense Review, 30-32.

Root, T. R., Hayes, J. F., Word, L. E., Shriver, E. L., & Griffin, G. R. (1979). Field test of techniques for tactical training of junior leaders in infantry units (Project EFFTRAIN) (ARI Technical Report 79-A21). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A075 604)

Sackett, P. R., & Larson, J. R. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed.). (Vol. 1., Section II, Chap. 8, pp. 419-489). Palo Alto, CA: Consulting Psychologists Press.

Scriven, M. (1967). The methodology of evaluation. In C. H. Weiss (Ed.), Evaluation action programs: Readings in social action and education (pp. 123-136). MA: Allyn & Bacon.

Siegel, A. I., & Wolf, J. J. (1969). Man-machine simulation models. New York: Wiley.

Sikora, J., & Coose, P. (1995). What in the world is ADS? Phalanx: The Bulletin of Military Operations Research, 28(2), 1-8.

Sottilare, R. (1995, March-April). More than a demo. . . Synthetic theater of war--Europe. Army RD&A, 18-20.

Sullivan, G. R. (1995). America's army: Strategic force for today and tomorrow. Defense, 1, 7-15.

Thompson, C. W. N., & Rath, G. J. (1974). The administrative experiment: A special case of field testing or evaluation. Human Factors, 16, 238-252.

U.S. Army Armor Center. (1993). Combat synchronization analysis: Final report. Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Army Armor Center. (1994, July). Advanced warfighting experiment: Operation Desert Hammer VI. Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Army Armor Center. (1995a). Battle lab experimentation plan for advanced warfighting experiment: Focused dispatch (Draft). Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Army Armor Center. (1995b). Force XXI: Creating the force of the 21st century [Slides]. Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-MW, Fort Knox, KY 40121-5200).

U.S. Army Armor Center. (1995b). Joint visitors bureau briefing [Slides]. Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-MW, Fort Knox, KY 40121-5200).

U.S. Army Armor School. (1995a, Feb.). Tactics, techniques and procedures for the digitized brigade (ST 71-3). Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Army Armor School. (1995b, Jan.) Tactics and techniques for the digitized battalion task force (ST 71-2-2). Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Army Armor School. (1995c, Jan.). Tactics and techniques for the digitized battalion task force (ST 71-1-1). Fort Knox, KY: Author. (Available from the Commander, U. S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200).

U.S. Department of the Army. (1994a, Dec.). Peace operations (FM 100-23). Washington, DC: Author.

U.S. Department of the Army 1994b, September). America's army--into the 21st Century. Washington, DC: Author.

U.S. Department of the Army. (1995a). Five year joint venture analysis plan (Draft Analysis Plan TRAC-JVAP-1, Annex III to Appendix C to Force XXI Campaign Plan). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1995b). Analytic support for the Army digitization office (ADO) axis (Draft Analytic Plan TRAC, Annex III to Appendix C to Force XXI Campaign Plan). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1995c). Version 1 Draft Warrior XXI Campaign Plan. Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1995d). Information operations (Field Manual 100-6). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1995e). Force XXI operations: A concept for the evolution of full-dimensional operations for the strategic army of the early twenty-first century (TRADOC Pamphlet 525-5, Draft).Fort Monroe, VA: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1995f, May). Battle Labs. . . defining the future. Fort Monroe, VA: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1996a). Land combat in the 21st century (Draft TRADOC Pamphlet). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1996b). Requirements determination (Draft TRADOC Pamphlet). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U.S. Department of the Army. (1996c). Training support package (TSP) development for collective and new equipment training (TRADOC Regulation 350-70). Washington, DC: Author. (Available from the Commander, U.S. Army Training and Doctrine Command, ATTN: ATDO-P, Fort Monroe VA 23651-5000).

U. S. General Accounting Office (1995). Battlefield automation: Army's digital battlefield plan lacks specific measurable goals (GAO/NSIAD-96-25). Washington, DC: Author.

Vowels, L. (1995, May-June). The lessons of Operation Desert Hammer VI: Materiel. Armor, 15-17, 34.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). Unobtrusive measures: Nonreactive research in the social sciences. Stokie, IL: Rand McNally.

Webster, W. G. (1995). What is the role of the CTCs in experimentation of the future force? In Combat Training Center Quarterly Bulletin, No. 95-11, pp. II-14 to II-18. (Available from the Center for Army Lessons Learned (CALL)< U.S. Combined Arms Command (CAC), Fort Leavenworth, KS 66027-7000).

Weiss, C. H. (1972). Evaluation research: Methods of assessing program effectiveness. Boston: Allyn & Bacon

Wheaton, G. R., Allen, T. W., Johnson, E. Drucker, E. H., Ford, P., Campbell, R. C., & Boycan, G. C. (1980). Methods for evaluating platoon battle run performance (ARI Technical Report 457). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A096 369)

Winsch, b. J., Garth, T. H., Ainslie, F. M., & Castleberry, J. D. (1996). Structured simulation-based training for digital forces: Initial battalion staff-level operations (ARI Research Product 96-08). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Witmer, B. G. (1981). A job aid for the structured observation of training (ARI Research Product 81-16). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A120 774)

Witsken, J. R. (1995, July-August). The lessons of Operation Desert Hammer VI: Our doctrine is basically sound. Armor, 35-37.

APPENDIX A

List of Acronyms

1SG . . . . . First Sergeant

AAR . . . . . After Action Review
ACTD . . . . Advanced Concept and Technology Demonstration
ASAS . . . . All Source Analysis System
ATD . . . . . Advanced Technology Demonstration
AWE . . . . . Advanced Warfighting Experiment

B2C2 . . . . Brigade and Below Command and Control
BCV . . . . . Battle Command Vehicle
BLEP . . . . Battle Lab Experimentation Plan
BLWE . . . . Battle Lab Warfighting Experiment
BMO . . . . . Battalion Maintenance Officer
BN . . . . . Battalion

C^2V . . . . . Command and Control Vehicle
CCTT . . . . Close Combat Tactical Trainer
CDR . . . . . Commander
CECOM . . . . Communications and Electronics Command
CO . . . . . Company
CTC . . . . . Combat Training Center
CTCP . . . . Combat Trains Command Post

DCE . . . . . Data Collection Exercise
DIS . . . . . Distributed Interactive Simulation
DTLOMS . . . Doctrine, Training, Leadership, Organization, Materiel, and Soldiers

ESP . . . . . Evaluation Support Package
EXFAC . . . . Experimental Facilitator

FA . . . . . Field Artillery
FD . . . . . Fire Direction
FIST . . . . Fire Integration Support Team
FTCP . . . . Field Trains Command Post

HV . . . . . Heavy

IFSAS . . . . Initial Fire Support Automated System
IPPT . . . . Integrated Product and Process Team
IVIS . . . . Intervehicular Information System

MBBL . . . . Mounted Battlespace Battle Lab
METT-T . . . Mission, Enemy, Terrain, Troops-Time
MEME . . . . Model-Experiment-Model-Evaluate
MEMV . . . . Model-Experiment-Model-Validate
MOE . . . . . Measure of Effectiveness
MOP . . . . . Measure of Performance
MORT/MTR . . Mortar
MWTB . . . . Mounted Warfare Test Bed

O/C . . . . . Observer and/or Controller
OPTEC . . . . Operational Test and Evaluation Command

PL . . . . . Platoon Leader

RDA&Tng . . . Research, Development, Acquisition & Training

SCT . . . . . Scout
SINCGARS . . Single Channel Ground to Air Radio System
STA . . . . . Station
STOW . . . . Synthetic Theater of War

TACOM . . . . Tank-Automotive Command
T&EO . . . . Training and Evaluation Outline
TF . . . . . Task Force
TOC . . . . . Tactical Operations Center
TRADOC . . . Training and Doctrine Command
TSP . . . . . Training Support Package
TTP . . . . . Tactics, Techniques, and Procedures

VTP . . . . . Virtual Training Program

XO . . . . . Executive Officer